DOCUMENT RESUME

ED 387 869 EA 027 057

AUTHOR	Burstein, Leigh; And Others
TITLE	Validating National Curriculum Indicators. Draft.
INSTITUTION	Rand Corp., Santa Monica, Calif.
SPONS AGENCY	National Science Foundation, Arlington, VA.
REPORT NO	DRU-1086-NSF
PUB DATE	Jun 95
NOTE	115p.
PUB TYPE	Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Data Collection; Data Interpretation; Elementary Secondary Education; *Evaluation Criteria; Evaluation Problems; *Instructional Materials; *Mathematics Curriculum; Mathematics Instruction; Mathematics Materials; *National Curriculum; *National Surveys; Research Methodology; *Validity
IDENTIFIERS	Educational Indicators

ABSTRACT

This report summarizes results from research aimed at improving the quality of information collected about school curriculum. The research sought to design and pilot a model for collecting benchmark data on school coursework. These more indepth data, such as course textbooks, assignments, exams, and teacher logs, can serve as anchors against which the validity of the survey items used in national data collections might be assessed. The data provide a basis for assessing the extent to which survey items measure what is taught in schools and classrooms. They can also be used to monitor whether the validity of teachers' responses have been undermined by outside factors. Data were derived from a survey of 70 mathematics teachers in 9 secondary schools located in California and Washington. The survey was administered before and after the collection of artifact data. Data were also collected from teacher daily logs, assignments, and interviews with principals, counselors, and mathematics department chairs. Chapter 2 details the study design, and the next three chapters summarize the extent to which major dimensions of curriculum can be measured through national surveys and then validated through deeper probes in a smaller number of sites. The final chapter discusses the implications for the design of future curriculum-indicator systems and for the policy uses of such information. It concludes that while an enhanced version of current national surveys can provide a reasonably accurate picture of high school mathematics teaching across the country, there are significant limitations on such data, and at this point, policy uses for more than informational purposes would be inappropriate. The study represents a first step in ensuring that curriculum indicators are valid and reliable measures of instruction. Nine figures, 12 tables, copies of the surveys, and a sample daily log form are included. (Contains 39 references.) (LMI)

ERIC

ED 387 869

D R A F T

RAND

Validating National Curriculum Indicators

Leigh Burstein, Lorraine M. McDonnell, Jeannette Van Winkle, Tor Ormseth, Jim Mirocha, and Gretchen Guitton

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

DRU-1086-NSF

June 1995

Prepared for National Science Foundation

The RAND restricted draft series is intended to transmit preliminary RAND research. Restricted drafts have not been formally reviewed or edited and have not been cleared for public release. Views or conclusions expressed in the drafts are tentative. A draft should not be cited or quoted without permission of the author.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

E. D. Gill

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

EA 027057

RAND is a nonprofit institution that seeks to improve public policy through research and analysis. RAND's publications do not necessarily reflect the opinions or policies of its research sponsors.

NOT CLEARED FOR OPEN PUBLICATION

ERIC

BEST COPY AVAILABLE

Chapter 1
INTRODUCTION

Efforts over the past decade to improve schooling have focused on changing the way schools are governed and on altering what they teach. The latter reform has been the more difficult to implement because it is multi-faceted--affecting curriculum content, teacher training, instructional approaches, and student assessment--and because classroom practice has traditionally been that aspect of schooling most insulated from the reach of public policy.

The multiple dimensions of curriculum and its varied manifestations in individual schools and classrooms have also meant that information about what students are being taught across the country is limited. This shortcoming has become evident as the demand for more comprehensive indicators, describing the status of U.S. schooling, has grown. A variety of policies such as the articulation of academic standards and new forms of student assessment assume that information is widely available on the content and modes of instruction. Yet most indicators of curriculum are limited to data collected by states on course offerings and enrollment, enumerated only by conventional course titles, and to national survey data based on student and teacher self-reports about course-taking, topic coverage, and instructional strategies. Both forms of data are inadequate. Many course titles convey no information about content or how that content is presented. Although the national data from sources such as the National Assessment of Educational Progress (NAEP) and the National Educational Longitudinal Study (NELS) are richer, no attempt has been made to determine whether information provided by teacher respondents is consistent with actual classroom practices and activities. Nor are there any explicit design features built into national indicator efforts to monitor whether responses are being corrupted by external events.

This report summarizes results from research aimed at improving the quality of information collected about school curriculum. Its purpose was to design and pilot a model for collecting benchmark data on school coursework. These more indepth data, such as course textbooks, assignments, exams, and teacher logs, can serve as anchors against which the validity of the survey items used in national data collections such as NAEP and NELS might be assessed. Together, these data constitute a series of deeper

ERIC

probes than are possible with survey data. As such, they provide a basis for assessing the extent to which survey items tap what is taught in schools and classrooms. They can also be used to monitor whether the validity of teachers' responses have been undermined by outside factors so that, for example, reports of classroom activities are consistent with current reform rhetoric, but are not matched by changes in actual practice. Benchmark data are more difficult and costly to collect, but they do not need to be collected as often or on as large a sample as conventional indicator data.

Although this research focused specifically on high school mathematics, many of its findings about measuring the multiple dimensions of curriculum also apply to other academic subjects taught in secondary schools. The study design is detailed in chapter two and the three subsequent chapters summarize the extent to which major dimensions of curriculum can be measured through national surveys and then validated through deeper probes in a smaller number of sites. A final chapter discusses the implications of our study for the design of future curriculum indicator systems and for the policy uses of such information. We conclude that while an enhanced version of current national surveys can provide a reasonably accurate picture of high school mathematics teaching across the country, there are significant limitations on such data and at this point, policy uses for more than informational purposes would be inappropriate.

Before turning to a description of our research methods, we provide some background for the study by discussing the research base on which it draws and the relevant policy and practice context.

RESEARCH BASE

A growing body of research documents the relationship between student achievement, the types of courses taken, and the content and level of those courses. Some of the most compelling evidence about the relationship between achievement and curricular content comes from the Second International Mathematics Study (SIMS), conducted between 1976 and 1982.[1] As their predecessors had recognized some twenty

[1]SIMS data collection included the administration of achievement tests and questionaires to over 125,000 students in 20 systems. Students were sampled from two population groups: Population A, consisting of students in the grade-level where the majority would be aged 13 and Population B, consisting of all students in the

ERIC

years earlier in the First International Mathematics Survey, the SIMS researchers understood that, in comparing student achievement across different national systems, curricular differences had to be taken into consideration. That recognition led to the notion of opportunity-to-learn (OTL). OTL became a measure of "whether or not...students have had an opportunity to study a particular topic or learn how to solve a particular type of problem presented by the test" (Husen as cited in Burstein, 1993: xxxiii). SIMS' researchers conceptualized the mathematics curriculum as functioning at three levels: the *intended curriculum* as articulated by officials at the system or national level; the *implemented curriculum* as interpreted by teachers in individual classrooms; and the *attained curriculum* as evidenced by student achievement on standardized tests and by their attitudes (Travers, 1993: 4).

The major vehicle for measuring the implemented curriculum was an OTL questionnaire administered to the teachers of tested students. Teachers were asked whether the content needed to respond to items on the achievement tests had been taught to their students. They were also asked more general questions about their instructional goals, their attitudes and beliefs about mathematics teaching, their instructional strategies, and their professional background.

One of the most important purposes that SIMS served was to document differences in curriculum, and hence in opportunities to learn, across national systems. For example, SIMS researchers found striking differences between the ways curricula are organized in the countries where students scored highest on the SIMS tests and the way they are organized in the United States. At the lower-secondary level, the Japanese curriculum emphasizes algebra; the curricula in France and Belgium are dominated by geometry and fractions. In contrast, U.S. schools allocate their curricula more equally across a variety of topics--thus covering each subject much more superficially. The mathematics curriculum in U.S. schools is characterized by extensive repetition and

---

terminal grade of secondary education who were also studying mathematics. In addition, supporting questionnaire data were collected from the approximately 6000 teachers of these students and 4000 principals or heads of mathematics departments. Information about the implemented curriculum was collected from National Committees which included mathematics educators and researchers (Travers, et al., 1988).

review, and little intensity of coverage. This low-intensity coverage means that individual topics are treated in only a few class periods, and concepts and topics are quite fragmented (McKnight, et al., 1987).

The SIMS data also illustrated variations in opportunities to learn within the same national system. For example, in the algebra content area, Japan's OTL ratings were quite similar across classrooms and teachers, ranging from 60 to 100 percent coverage of the content included in the SIMS test items, with the median topic coverage 85 percent. In contrast, the United States' OTL ratings ranged from 0 to 100 percent, with the median at 75 percent of the SIMS topics. About 10 percent of the thirteen-year-olds tested in the U.S. were receiving virtually no instruction in algebra topics and 25 percent were receiving instruction in only about half the mathematics content covered on the algebra sub-test (Schmidt, et al., 1993).[2] Part of the reason for the greater variation in OTL in the United States is that schools typically assign students to different kinds of mathematics classes according to their abilities. Using the SIMS data, Kifer (1993) found that when the U.S. eighth grade mathematics classes in the SIMS sample were categorized as remedial, regular, enriched, and algebra, significant differences were evident in students' opportunities to learn the content tested in SIMS. With the exception of arithmetic topics, students in remedial classes receive very little teaching on mathematics content, and even students in regular classes receive less content coverage in algebra and geometry topics than those in enriched and algebra classes.

The SIMS results visibly influenced public discussion because they showed significant gaps in U.S. students' achievement, as compared with students in other industrialized countries. But other studies focused solely on the United States produced similar findings about the effects of curricular exposure. For example, Raizen and Jones (1985) summarized four studies based on nationally representative student samples that showed a strong correlation between the number of mathematics courses students take and their achievement in mathematics. These relationships persist even when

[2]On the whole, opportunities-to learn were considerably more uniform in France and Japan than in the United States and New Zealand, although within-system variation is greater for all systems in geometry than for either algebra or arithmetic topics (Schmidt, et al., 1993).

ERIC

background variables such as home and community environment and previous mathematics learning are taken into account. Research had also shown that the level, as well as the number of courses students take is correlated with achievement. Jones et al. (1986), after controlling for socioeconomic status and test scores two years earlier, found that students in the High School and Beyond (HS&B) sample with at least five transcript credits in mathematics at or above the algebra I level scored an average of 17 percentage points higher on a standardized mathematics test than those with no course credits in higher-level mathematics. In documenting that curricular exposure is a significant predictor of student achievement and a critical factor in influencing the distribution of students' learning opportunities, all these studies make a strong case for supplementing data on student achievement with information about the curriculum they experience.

POLICY AND PRACTICE CONTEXT

Growing concern about the achievement of U.S. students and the distribution of that achievement across different types of students has also prompted an intensified focus on school curriculum as a focal point for policy interventions. Beginning in the mid-1980s, elected officials, especially at the state-level, extended their traditional concern about how schools are governed and financed to include what schools teach, who teaches it, and in some cases, how it is taught. In fashioning policies in this area, policymakers drew on the research that demonstrated the close link between students' curricular exposure and their achievement, and on expert advice about what constitutes an engaging, productive curriculum.

Recent examples of this focus are the federal Goals 2000 legislation and similar standards-setting exercises in the states. Goals 2000 provides grants to states as inducements for them to establish curriculum and student performance standards, as well as standards or strategies that ensure students will have an opportunity to learn the content embodied in the state standards. Even prior to the federal effort, however, a number of states were already using curriculum as a reform vehicle by relying on such strategies as the development of curricular standards and frameworks, the redesign of their assessment systems, and other means such as textbook adoption policies.

These federal and state initiatives have drawn on the prior, standards-setting efforts undertaken by professional organizations and have also prompted other disciplines to begin similar exercises. The National Council of Teachers of Mathematics (NCTM, 1989, 1991) was responsible for one of the earliest professional efforts to improve classroom practice through the promulgation of curricular and teaching standards. Its approach was later reflected in new state curriculum frameworks such as those in California (California Department of Education, 1992). In mathematics, curriculum reform has been characterized by learning goals that emphasize understanding the conceptual basis of mathematics, reasoning mathematically and applying that reasoning in everyday situations, offering alternative solutions to problems, and communicating about mathematical concepts in meaningful and useful ways. Consistent with those goals, curriculum reformers have advocated changes in both mathematics content and instructional strategies. Particularly prominent in this reform vision of the mathematics curriculum is a changed view of the teacher's role. Because students are expected to play an active part in constructing and applying mathematical ideas, teachers are to be facilitators of learning rather than imparters of information. In terms of actual instructional activities, this shift means that rather than lecturing and relying on a textbook, teachers are to select and structure mathematical tasks that allow students to learn through discussion, group activities, and other modes of discovery.

Despite its growing popularity, the use of curriculum as a lever for educational reform is not without its problems. Most of the attention thus far has focused on the political difficulties inherent in defining what should be included in state curriculum standards. The recent experience of states like California, Kentucky, and Pennsylvania where serious controversies have erupted over the content of state curriculum frameworks and assessments, illustrate the passion that questions about what students should be taught can engender (Merl, 1994; Harp, 1994; Ravitch, 1995). Debate has also erupted over the use of OTL standards as part of a curricular reform strategy, with the controversy focused on values such as how equity is defined or the appropriate role of state vs. local government (National Council on Education Standards and Testing, 1992; O'Day and Smith, 1993; Rothman, 1993; Owens, 1994; Goodling, 1994).

Equally important, however, are the technical feasibility issues that arise when curriculum is used as the focus of policy. One major problem stems from limitations on the amount and type of indicator data currently collected by the federal government and the states. Statistical data about the condition of schooling focused historically on inputs such as per pupil spending and on outcomes, most notably student test scores. Information about how schools are organized and how students are taught tended to be available only through research studies that were based on data collected from limited samples on a non-routine basis.

However, beginning in the mid-1980s, a number of researchers and policymakers began to advocate expanding the type of indicator data that was routinely collected and reported (Murnane and Raizen, 1988; Shavelson, et al., 1987; 1989; OERI State Accountability Study Group, 1988; National Study Panel on Education Indicators, 1991; Porter 1991). They argued that indicator data on school and classroom processes were necessary to monitor educational trends, compare schooling conditions across different kinds of students in different geographic locations, and to generate information that could be used in holding schools accountable. A good part of the rationale for collecting more than just input and outcome data lie in the fact that these indicators were to be used for policy purposes. Knowing that educational conditions were getting better or worse provided little insight into why particular trends existed or how to fix problems or replicate successes. Furthermore, it had become clear that the way in which educational inputs were used was as important as the absolute level of those resources. To accommodate the information needs of policymakers, then, indicator systems had to include data that could provide a comprehensive picture of the schooling process as it occurred in schools and classrooms.

Consequently, proposed designs for new indicator systems advocated including process measures such as teacher background and experience, school- and grade-level organization, course offerings and student course-taking patterns, curriculum content, instructional materials availability and usage, and instructional strategies. In recommending that a broad array of school and classroom process measures be included in indicator systems, researchers drew upon studies documenting the relationship between student achievement and the type of instruction they receive (Shavelson, et al., 1989).

Some indicator systems were expanded to include school process data. For example, at the national level, NAEP and the longitudinal surveys of students sponsored by the National Center for Education Statistics (NCES) (e.g., HS&B, the NELS) surveyed students, teachers, and school administrators about school organization and resources, teacher qualifications, curricular content, and instructional strategies. These data could be disaggregated by gender, ethnicity, urbanicity, and in some cases, by state. In addition, in mathematics and science, 47 states were reporting data to the Council of Chief State School Officers on teacher qualifications and student course-taking patterns (Blank and Gruebel, 1993). By the late 1980s, in about half the states, data were available about school-level performance (e.g., student test scores, attendance, and drop-out rates) and some states like California issued "school report cards" which included school process data such as the proportion of students taking college preparatory or Advanced Placement courses (OERI State Accountability Study Group, 1988). Although not typically reported by school, many states also collect information about teacher qualifications that could be disaggregated to the school-level. No one, however, is collecting data on the curricular content and instructional strategies available to students in different local jurisdictions. At this point, it is possible to describe how students' curricular opportunities differ for boys vs. girls, for different ethnic groups, and for urban students as compared with those in either rural or suburban areas. But we do not know whether the curriculum experienced by students in Seattle is significantly different than what students in Indianapolis or Pikeville, Kentucky experience, or whether curriculum differs greatly among schools within the same state.

Besides these limits on the amount and type of curriculum indicator data, there are substantial methodological problems with the available data. The most common data, available on a school-by-school basis, is derived from reports by principals and other administrators about course offerings and student enrollment in those courses. However, the SIMS data suggest that because of significant variation in the breadth and depth of topic coverage, knowing that most ninth graders take algebra does not provide adequate information about their actual opportunity to learn algebra content.

Even the more comprehensive data about classroom processes, collected from nationally-representative samples of teachers, are limited in their ability to portray a valid picture of the schooling process. Most curriculum data are collected through teacher surveys because these are cost-effective and impose only a modest time burden on respondents. However, some aspects of curricular practice simply cannot be measured without actually going into the classroom and observing the interactions between a teacher and students. These include: discourse practices that evidence the extent of student participation and their role in the learning process, the use of small group work, and the relative emphasis placed on different topics within a lesson and the coherence of teachers' presentations. Given the rudimentary status of curriculum data in most national and state indicator systems, efforts to obtain an accurate picture of how opportunities to learn vary for different groups of students will most likely continue to focus at a more general level than these finely-grained aspects of instruction.

If policymakers and the public are interested in data about school curriculum that are both comparable across local jurisdictions and can be disaggregated to the school-level, teacher surveys will remain the most feasible way to collect such information for the foreseeable future. Yet, to this point, none of the national survey data collected from teachers has been validated to determine whether it measures what is actually occurring in classrooms. Despite major advances in the design of background and school process measures, studies have generally developed a few new items and then "borrowed" others from earlier studies. Little effort has been made to validate these measures by comparing the information they generate with that obtained through alternative measures and data collection procedures. For example, are teachers' reports of curricular goals or content coverage consistent with the material tested and the types of questions asked on their exams?

Given the complexity of the teaching and learning process; the amount of variation across classrooms, as evidenced from more indepth, school-based research; and shifting modes of instruction as new curricular reforms are introduced, it is reasonable to assume that surveys alone may not adequately measure even the most generic forms of instructional practice. Therefore, if national teacher surveys are to remain the major source of information about the instruction American students are receiving and if policy decisions continue to be made based on these data, mechanisms will need to be established to validate the survey data. The benchmarking strategy which relies on other data such as textbooks and teacher assignments, that is outlined in this study, is one method for improving the quality of national curriculum indicator data.

Past research on the relationship between student achievement and the instruction they receive, as well as the growing emphasis on curriculum as a policy lever, suggest several factors that need to be considered in efforts to improve the quality of curriculum indicator data--whether they be based solely on surveys or also include more indepth validation procedures. First, curriculum is a multi-dimensional concept that includes, but is not limited, to the content of instruction. Consequently, in addition to content or topic coverage, information also needs to be collected on several other dimensions. An obvious one is teachers' instructional strategies. Key elements include the manner in which content is sequenced and the mode in which teachers and textbooks present it to students. For example, the effect on student learning might be quite different if a teacher presents new content through a lecture than if she introduces students to the same content by asking them to apply previously-learned concepts to a new situation and has them do it while working in small groups. Another critical dimension of curriculum are the goals that teachers pursue as they present course content to students and use various instructional strategies. The relative emphasis that teachers give to different objectives reveals something about their expectations for a particular course, and their choice of objectives is likely to influence how they configure topics and instructional activities within that course. However, teachers' reports of their course objectives reflect intended behavior and are likely to be less reliable than reports of actual behavior, such as topic coverage and instructional activities. For that reason, data on teacher goals can be suggestive, but they need to be interpreted in tandem with other information about classroom activities.

Second, curriculum indicators need to capture the variability inherent in a complex activity such as teaching. We have noted that data on course enrollments alone are insufficient because they convey little information about the actual content of the course and even less about the instructional strategies used. Similarly, because of the current flux in instructional policy and practice, data collection instruments need to measure both traditional forms of instruction and the newer approaches advocated by curriculum reformers. Strategies such as having students work in small groups to find joint solutions or use manipulatives to demonstrate a concept are currently much in vogue among reformers. Yet decades of research on educational change, and most recently on the implementation of curriculum reforms (e.g., Cohen and Peterson, 1990), suggests that many teachers will continue to use more traditional approaches such as lecturing to their students and having them work exercises from a textbook. Therefore, data collection instruments need to be broadly-focused and sensitive enough to reflect the diversity of classroom practice during a transitional period in school curriculum.

In the next chapter, we outline our study methods and indicate how we attempted to take past research and the current context into consideration in designing this study.

Chapter 2
STUDY METHODS

GENERAL APPROACH AND RATIONALE

The benchmarking procedures developed in this study were designed as one way to validate survey data collected from classroom teachers. However, a number of different approaches could be used to ensure that survey items accurately measure what is happening in classrooms. In choosing among possible strategies, two criteria need to be considered. Any validation strategy should measure curricular goals, content, and instructional activities as sensitively as possible, but it must also do so cost-efficiently without imposing a significant burden on teachers and students.

The methodology of teaching and learning research would suggest that detailed classroom observations are the best way to make inferences about the curriculum students are actually receiving (for a detailed discussion of various approaches to teaching and learning research, see Wittrock, 1986). However, from a national indicators perspective, this approach is problematic.

Although classroom observation is an effective method for capturing curricular depth, it is considerably less efficient in measuring breadth--a requirement for indicator purposes. For example, if one's purpose were to focus intently on a narrow slice of curriculum (e.g., the teaching of the Pythagorean Theorem) taught at a prescribed point in most classrooms of a given course, then one could target a specific amount of observation time to capture the teaching of that topic, and comparing survey responses with observational data would presumably be straightforward. But for most purposes, the span of curriculum to be measured through indicator data is much more extensive, and the sequencing of topics and time allocations vary considerably from section to section of even the same course, much less across courses. It may well be that instruction on certain topics cycles throughout a course, making the targeting of observation even more impractical. Choosing a fixed time of the school year to conduct observations and capture whatever topics might be taught at that time runs the risk of misspecifying the place within a specific teacher's curriculum the observed topic falls, and missing what was covered previously and planned for later. Consequently, the only kinds

ERIC

of survey questions that could be validated in this way would be general ones dealing with activities and process, as distinct from content.

Such limitations led us to conclude that on cost and feasibility grounds alone, classroom observation was not a viable tool for obtaining ongoing benchmark data. While it is an appropriate and necessary strategy for basic research on school curriculum, classroom observation is not practical for education indicator purposes. Moreover, unless observations were long-term and extensive, they could very well distort decisions about the validity of specific survey alternatives.

Consequently, we decided to build on prior research (McDonnell et al., 1990) and to make the collection and analysis of a representative sample of teacher assignments (homework, quizzes, classroom exercises, projects, examinations), gathered throughout the semester, the centerpiece of the benchmarking effort. We believe that these examples of classwork and how the teacher uses them represent much of the curriculum as experienced by students. Thus these systematic artifacts of learning, placed within the context of syllabi and textbook coverage, constitute a solid basis for characterizing the implemented curriculum presented to students.[3] In addition, by spreading data collection over a broader period of time, at a much lower cost than equivalent observational activities, the span of curriculum that can be measured is expanded considerably.

The approach taken in this study, then, was to use these instructional artifacts as deeper probes about the nature of instruction in a small number of sites. The artifacts were coded to extract data about teachers' instructional content, activities, and goals. That information was then compared with their responses on surveys similar to those

---

[3] However, these artifacts do not provide information about how students receive and respond to the curriculum, only how teachers present it. In our pre-test, we asked teachers, for each major assignment provided, to include two samples of student work graded as an A, two B/Cs and two examples of below-C work. However, that request created an extra burden for respondents (even when arrangements were made to have the student work copied for teachers), and most of the non-response rate for the study was accounted for by this request. Therefore, we did not request student work from the remainder of the teachers in the sample.

In the future, requests for student work may become less burdensome and intrusive as more schools adopt student portfolios, and routines are established for the systematic production, copying, and storage of student work.

BEST COPY AVAILABLE

administered as part of national data collection efforts.[4] The overarching question was whether measures of goals, activities, and content from the survey cohered or were correlated with similar measures obtained from the benchmark data. To the extent that inconsistencies emerged, we needed to analyze why and to identify ways to improve coherence in future indicator data. The results of this endeavor are threefold: an analysis of how well survey data measure curriculum, as compared with data that are closer to the actual instructional process; a recommended set of procedures for periodically validating data collected from large-scale surveys; and suggested enhancements in the type and number of items included on these surveys.

STUDY SAMPLE

The study is based on data collected from a sample of 70 teachers who comprise the majority of the mathematics faculty in nine secondary schools located in California and Washington.[5] The characteristics of the schools and the teachers are summarized in Table 2.1. Although the indepth and exploratory nature of the data collection meant that only a small sample of teachers could be studied, we wanted to make certain that they were typical of those who participate in large, national surveys. Therefore, schools were selected from among those that were part of the 1992 NELS Second Follow-up Study (NELS-SFU).[6] Twenty-four schools were contacted and nine agreed to participate.

---

[4]The artifacts were coded by six experienced mathematics teachers and two project staff, using a coding instrument that paralleled the items on the survey. The coding process is described in a subsequent section.

[5]In addition to the mathematics teachers from whom data were collected, 18 science teachers from seven of the sampled schools also participated in the study as part of an exploratory analysis focused on developing curriculum indicators for high school science courses. However, this report is based on only the data collected from the mathematics teachers.

Each participating teacher was paid an honorium of $175 to complete two surveys and provide instructional artifacts over the course of a semester. The 13 teachers who participated in follow-up interviews were paid an additional $50.

[6]Because NELS was designed to obtain data on a nationally-representative sample of students, teachers were included only if they taught students in that sample. Therefore, the 2606 mathematics teachers who were surveyed in NELS-SFU do not constitute a nationally-representative smaple of high school mathematics teachers. However, just to show how our much smaller sample compares with a larger one drawn from across the country, we compared our teachers with the NELS-SFU sample and found that the mean years of teaching experience is exactly the same for the two groups. Our sample has a slightly higher proportion of males (58 percent as compared with 52 percent for NELS-SFU), but the major difference between the two groups is that our sample

ERIC

Table 2.1
STUDY SAMPLE

School Characteristics	(N=9)	Number
California		
Urban		4
Suburban		1
Rural		1
Washington		
Urban		1
Suburban		2
Mathematics classes in each of the course categories examined:		
Below Algebra I		20
Algebra I		15
Geometry		12
Algebra II/Trigonometry		8
Math Analysis/Pre-Calculus		7
Calculus		8

Teacher Characteristics	(N=70)*	Percent
Male		58
Female		42
College major in mathematics		47
Mean years of teaching experience	17 (S.D.=9)	

*74 teachers agreed to participate in the study, but four dropped-out before the artifact data collection was completed.

ERIC

Of the remaining, nine refused to participate. The others agreed to participate, but they were eliminated for various reasons such as the small size of the mathematics faculty in several schools and year-round schedules that did not coincide with our data collection timetable.

Of the nine schools, five are located in urban areas, three are suburban, and one is rural. The largest school enrolls 2800 students, but five schools have enrollments in excess of 2000. The smallest school enrolls 980 students. The enrollment in five of the schools is 65 percent or more Anglo, while the other four have minority enrollments of 65 percent or more.

DATA

Table 2.2 summarizes the types of data collected and the purpose each data source served in the study design. All the data are discussed at greater length in this section.

Teacher Surveys

Three factors shaped the design and administration of the survey component of the study. First, because the purpose of the project was to validate data collected as part of efforts such as NELS, the survey instrument needed to approximate closely the type administered in national surveys. Second, we had to make certain that the collection of artifact data did not bias teachers' sur responses by sensitizing them to the kinds of questions that would later be asked of them on the survey.[7] Finally, we wanted to pilot the administration of a more extensive survey than has typically been used in national indicator data collection.

---

includes a considerably lower proportion of teachers with a college major in mathematics (47 percent as compared with 70 percent in the NELS-SFU sample).

[7] Our concerns about artifact data collection contaminating survey responses were two-fold. The first was that if teachers were completing daily logs and providing assignments throughout the semester, they might become more aware of the types and frequency of their classroom activities than they would ordinarily be. Consequently, their survey responses would be more accurate than would be the case in routine data collection when teachers only complete a survey. If that were the case, the survey responses in our study would not be equivalent to those collected in national indicator efforts. Second, we were concerned that because of their direct contact with members of the research team throughout the semester, teachers might be more likely to give what they considered to be socially desirable responses. In this case, those responses were likely to be consistent with the rhetoric of the mathematics reform movement and away from more traditional teaching strategies.

Table 2.2
STUDY DATA

Type of Data	Purpose
Survey (administered *prior* to artifact data collection)	• Obtain teacher self-reports about topic coverage, instructional strategies, and goals on a survey instrument analogous to the format and level of detail on national surveys. Data also collected on teacher background and experience. • Serve as a basis for gauging the extent to which teachers' responses were altered by the artifact data collection.
Survey (administered *after* artifact data collection)*	• Obtain more extensive teacher self-reports with a more detailed list of topics, instructional activities, and goals and with items that tap these dimensions in multiple ways.
Course Textbooks	• Validate survey data on topic coverage
Teacher Daily Logs (5 weeks)	• Obtain more precise self-reports from teachers on topic coverage and instructional activities.
Daily Assignments (5 weeks) Exams and Quizzes (5 weeks) Major Assignments and Projects (entire semester)	• Validate survey data on topic coverage, goals, and instructional strategies (viz., assignment/exam format and characteristics)
Interviews with principals, counselors, and mathematics department chairs	• Obtain information on school context -- student characteristics, the different levels of courses offered, and how teachers and students are assigned to courses
Follow-up group interviews**	• Identify reasons for anomalous findings

*86% of the sample completed both the pre- and post-data collection surveys. The remainder completed one of the two surveys.

**Thirteen teachers from two of the sample schools participated in the follow-up group interviews.

Our strategy for taking these factors into consideration was to administer a survey prior to collecting the artifact data. It included the same items as those in the instructional activities, content, goals, and teacher background sections of the teacher questionnaire administered as part of the 1992 NELS-SFU. Teachers were asked to respond in terms of one particular section of a single course that they were teaching. We then collected artifact data on that same section over one semester. At the end of the semester, we administered a second survey that repeated the same instructional activities and content items asked on the first survey, but also included an expanded list of topics, goals, and instructional activities. The NELS-SFU survey contained 11 topics to measure content coverage, 16 items on instructional strategies, and ten on goals and objectives. In contrast, the survey administered after the artifact data collection included 30 topics to measure content coverage (with separate topic lists for courses at or above algebra II and another for courses below that level), 33 items on instructional strategies, and 32 on goals and objectives. The enhanced survey also included items designed to measure teachers' expectations about levels of student understanding and how teachers conceive of their role in student learning. Appendix A contains copies of both questionnaires.

In addition to expanding the post-data collection survey to probe in greater depth and to measure curriculum in more diverse ways, we also experimented with a variety of different item formats and response options. For example, the NELS survey asks teachers whether a topic was *taught previously, reviewed only, taught as new content, will be taught or reviewed later in the year*, or whether the *topic is beyond the scope of the course or not included in the curriculum*. In addition to this response option, the enhanced survey also asked about the number of periods spent on a topic, using a response option that included six categories ranging from *0 periods* to > *20 periods*. In some questions, teachers were asked to describe characteristics of their instructional activities in terms of the percentage of class time or of an assignment; responses were elicited in some cases as a continuous variable and in others, as a categorical variable. In other questions, frequency was defined as a categorical measure ranging from *almost* *every day* to *never*. Similarly, teachers were asked about the amount of emphasis they give to different goals, but they were also asked more indirectly about curricular goals in a question that probed their expectations for students' level of understanding. Including a variety of different types of response options provided us with another source of information from which to make recommendations about how to improve existing surveys.

Analysis of the two surveys suggests that teachers' responses were not biased by the artifact data collection, and that validation procedures can be designed to occur after survey data have been collected. When we compared teachers' responses to the two surveys, we found few significant differences between their responses on items that appeared on both the pre- and post-survey. On average across all items common to both surveys, 90 percent differed by no more than one response option and 60 percent were exactly the same on the two surveys. Those items where a large proportion of responses changed were ones that would be expected to change between the beginning and end of the semester because teachers have more precise information at the end--e.g., the percent of class time spent administering tests or quizzes, the frequency of teacher-led discussions. In addition, there was no evidence that teachers gave socially desirable responses, or felt it necessary to present an image of their teaching consistent with the rhetoric of the mathematics reform movement. As the discussion in the subsequent chapters will indicate, a large proportion of teachers reported engaging in traditional activities such as lecturing and correcting or reviewing homework on a daily basis, and most reported engaging in reform-oriented activities such as student-led discussions rarely or not at all.

Instructional Artifacts

Course textbooks. A copy of the textbook used by each teacher in the study sample was purchased, and teachers were asked on the post-data collection survey which chapters they had covered over the course of the semester and which additional ones they had already or planned to cover during the rest of the year.[8] All the chapters or lessons a teacher reported as covering were then coded to determine which topics were covered. That information became one of the benchmarks against which topic coverage, as reported by teachers on the survey, was compared.

Teacher daily logs. During the same five weeks that all their daily assignments were collected, teachers were also asked to complete a one-page log form (included in Appendix A) at the end of each day. The form asked them to list which topics they covered during that day's class period and to indicate on a checklist all the modes of instruction they used and the activities in which the students engaged. There was also a *comments* section where teachers were asked to provide any information about the lesson that they felt was important (e.g., that class time was reduced by other school activities, that something particularly different or innovative occurred that day). In order to minimize teacher burden, the log form was designed to be completed in approximately five minutes.

Because the logs were completed by teachers, they do not represent an external source for validating the surveys in the same way that textbooks and assignments do. However, they do provide a check on the reliability of the surveys since they provide greater detail about classroom activities, with the information collected closer in time to the actual events.

Assignments. Teachers were asked to provide copies of every assignment they gave to students for a period of five weeks. The five weeks of data collection were divided into one week at the beginning of the semester, three consecutive weeks in the middle, and one week at the end. During these times, teachers provided all in-class and homework assignments, quizzes, exams, major projects, and any other written work assigned to students. In addition, teachers completed a pre-printed label, checking the

[8]Four teachers did not use a textbook. Two teach interactive mathematics which is an alternative method for teaching algebra and geometry that combines the two subjects and integrates individual topics within a problem-solving focus. The other two teach Math A-B which is a course offered in California schools for those students who need to take a preparatory course prior to beginning algebra. One other teacher in the sample used a textbook published more than ten years ago that is now out of print. Consequently, this data source was not available for five teachers.

major purpose of an assignment, its relationship to other classwork, whether the work was done individually or in groups, and whether done inside or outside the classroom. This label was affixed to each assignment. During the remaining weeks of the semester, teachers provided copies of their major assignments only--i.e., exams, papers of more than three pages, and projects. A pre-printed label was also attached to each of these assignments. On average, 20 assignments, including major assignments and projects, were provided by each participating teacher (n=1407). Exams and quizzes averaged about five per teacher (n=368).

Interviews

In each school, we conducted face-to-face interviews with the principal, the head counselor, and the mathematics department chair. These interviews averaged about 45 minutes, and focused on the type of students attending the school, the different levels of courses offered, what criteria the school used in assigning students to different mathematics courses and sections, and how decisions about teacher assignments were made. We also asked the department chairs to describe in some detail the major differences among the mathematics courses offered by the school in terms of level of difficulty, types of students enrolled, topics covered, instructional materials and strategies, course requirements, and grading practices. These interviews helped us place the survey and artifact data in a richer and more valid context. We were particularly interested in finding out whether there were any recent school- or department-level initiatives that might be shaping the curricular content or instructional approaches used by teachers.

We had not planned to conduct any follow-up interviews with teachers after the artifact data were collected. However, we were having difficulty interpreting several key findings that showed a lack of internal consistency between what teachers reported on the survey as their goals and what they reported about instructional activities. We found, for example, that a substantial proportion (40 percent) of teachers were reporting a *major* or *moderate* emphasis on most of the goals consistent with the mathematics reform movement. However, only a small proportion (12 percent) reported engaging regularly in most of the instructional activities advocated by NCTM. Similarly, the mean level of agreement between teachers' self-reports about their goals on the surveys and the coding

ERIC

of their exams was low. The typical pattern was for teachers to report a *minor* or *moderate* emphasis on most goals, while coders judged teachers' exams as showing *no* emphasis on those goals. The discrepancy was greatest on the so-called "reform" goals and considerably less on more traditional goals (e.g., *performing calculations with speed and accuracy*).

Before we concluded that these discrepancies represented "real" differences between teachers' reported and actual behaviors, we wanted to make certain that they were not the result of fundamentally different understandings between teachers' interpretation of survey items and the coders who were using the reform movement's definitions. As a result, we decided to address these questions through the use of follow-up, group interviews. We interviewed all the original study participants from two high schools in several group discussions that lasted about 90 minutes each. We asked teachers questions that would help us clarify our anomalous results. For example, with regard to the instructional goals that seemed to have been interpreted inconsistently, we asked: "in the course you reported on in your survey, what types of instructional activities do you see as representing this particular goal?" We report the results of these group interviews in subsequent chapters as one basis for interpreting some of our findings.

The study data were collected in four waves. We initially collected data from teachers in two schools in the spring of 1992, as a pilot for the rest of the study. We found no substantive problems with our data collection instruments and procedures, but we needed to streamline them to reduce teacher burden. Consequently, the request for graded student work was eliminated and the enhanced teacher survey was shortened. We also wanted to make certain that there would be no significant differences between collecting data in the fall, as compared with the spring semester. Consequently, we collected data from three additional schools in fall 1992 and from the remaining four in spring 1993. The follow-up group interviews were conducted in March 1994.

CODING THE ARTIFACT DATA

The effectiveness of a validation strategy, based on instructional artifacts, rests entirely on how information is coded or extracted from those artifacts. Valid and reliable coding requires that three criteria be met. First, in order to make comparisons

ERIC

between the survey and the artifacts, the coding format needs to parallel the survey items as closely as possible. However, valid comparisons depend on more than just a similar format for the two types of data. The survey items and the coding categories should be so clearly defined that teachers and coders will interpret them similarly. Second, the coding should extract as much information as possible from the artifacts so as to provide a full, valid description of a teacher's instruction, but it needs to do so without requiring judgments or inferences that go beyond the data. Third, the data need to be coded reliably--i.e., another coder would make similar judgments about the same information.

Several factors work against meeting these criteria, however. First, artifact data are unstandardized in the sense that the type and mix of assignments can vary considerably across teachers. Even the textbooks in our sample, the most standardized type of artifact, varied from the conventional (e.g., Dolciani's Algebra I text) to the innovative (e.g., Sunburst Geometry, Merrill's integrated math series) to the controversial (the Saxon series). Second, while some dimensions of curriculum have commonly-understood meanings, others do not. For example, most mathematics teachers would agree on what content falls within the categories of *square roots*, *quadratic equations*, or *slope*. But topics such as *math modeling* or *proportional reasoning* may be interpreted quite differently by different teachers. As we found in our analysis, the problem is particularly acute for curricular terms associated with the mathematics reform movement.

Third, coding a given teacher's artifacts requires a large number of judgments, some of which may require inferences that go beyond the available data. Although textbooks only need to be coded for topic coverage, other artifacts have to be coded to extract information on topics, instructional characteristics of the exam or assignment, level of understanding required of students, and teachers' instructional goals. Depending on the degree of aggregation desired, coding judgments can be made across all artifacts of a given type (e.g., across all assignments); with each separate exam or assignment as the unit of analysis; or at the most disaggregated level, on an item-by-item basis within a given assignment or exam.[9] In addition to the sheer number of judgments, coding artifacts also requires a variety of different kinds of judgments. In some cases, it only involves matching a textbook lesson, assignment, or exam item to one of the topics on the survey list. But other coding tasks require more complex judgments--e.g., identifying types of exam and assignment formats, making inferences about the purpose of an assignment or about a teacher's instructional goals. The number and variety of judgments involved in coding a teacher's artifacts provide considerable detail about the nature of his or her instruction, and expand the number of benchmarks available to validate the survey results. The downside is that the greater the number and the variety of judgments that coders have to make, the more difficult it is to ensure an adequate level of reliability.

We addressed these constraints on valid and reliable coding by using six experienced, secondary mathematics teachers as coders. Project staff trained them for two days and the coders were then supervised by two project staff who are also experienced mathematics teachers. During the two days, coders familiarized themselves with the coding manual and sample sets of artifacts. They also did practice coding, followed by an extended discussion and refinement of the coding rules. The first artifact file took each coder about one day (approximately 7 hours) to complete, but the amount of time was reduced to about 2-4 hours per file once coders became more experienced.

About 15 percent (n=11) of the artifact files were double-coded by project staff for reliability purposes. The rate of consistency between coders varied somewhat across the types of artifacts. For textbooks, coders had a rate of agreement of 58 percent on the exact number of lessons that included a particular topic, 74 percent of their judgments about topic inclusion differed by only one lesson, and 85 percent were within

[9] In our coding, we chose an approach that falls somewhere in the middle of these three options. Topic coverage, level of understanding, and assignment characteristics were coded for assignments (homework, in-class exercises, quizzes) at the level of the individual assignment. However, coders were asked to make summary judgments about teachers' goals as they were evidenced across all their assignments (i.e., one judgment based on their approximately 20 assignments). For exams, the coding was done at a finer level of detail, with level of understanding coded for each individual item or question on an exam, and instructional goals for each separate exam. Closer attention was paid to exams because we felt that while both assignments and exams represent the enacted curriculum, exams communicate what teachers consider to be most important.

ERIC

two lessons. On assignments, the coders had a rate of agreement of 74 percent on all their judgments about topic inclusion, instructional characteristics, and goals; they differed by only one category for 86 percent of the judgments they made, and were within two categories on 91 percent. The rates for exams were 71, 78, and 81 percent respectively for the three levels of agreement.

Although these rates of agreement are reasonable, given the nature of the task, two caveats are in order. First, these aggregate rates of agreement mask the large number of judgments that coders had to make. For example, for each assignment, coders were making 30 different judgments; a number then multiplied across the approximately 20 assignments each teacher provided. For exams, the number of separate judgments was 51, multiplied by the 5 or so exams from each teacher. A second caveat points to what became an important factor in interpreting some of our substantive results--viz., that the aggregated rates mask considerable variation across types of judgments. On some items, the rates of agreement between coders were close to 100 percent and in other instances, they fell below 50 percent. The items with the highest rates of agreement tended to be the more specific, narrower content topics (e.g., *complex numbers*) and traditional instructional approaches and goals (e.g., proportion of exam items that are multiple choice, proportion that are minor variations of homework problems). Those with the least agreement were either broad topic categories or more reform-oriented topics and approaches (e.g., *patterns and functions, problems [having] more than one possible answer*). Although our coders were experienced teachers, conversant with the NCTM standards, and trained in a common set of decision rules, their lack of agreement evidenced some of the same confusion about terms that was reflected in teachers' responses. As a result, these coding problems helped inform our substantive findings and recommendations for improving future data collection.

In the next three chapters, we summarize major findings, focusing first on instructional content, then instructional strategies and finally, on instructional goals. In each chapter, we provide examples of the kinds of information about curriculum that can be obtained from teacher surveys. We then examine the level of consistency between survey responses and the artifacts, identify reasons for discrepancies, and suggest how they might be reduced in future indicator efforts.

Chapter 3

INSTRUCTIONAL CONTENT

Instructional content, or the topics covered in a particular course, form the core of the implemented curriculum. Although it is mediated through the instructional strategies that teachers use, content is the dimension of curriculum whose relationship to student achievement is the most well-established. It is also the aspect of curriculum that has proven the least problematic to measure through teacher surveys. National surveys such as NAEP and NELS have typically asked teachers whether they taught or reviewed any of the items on a general list of topics. By asking teachers whether their students had been taught the content reflected in specific test items, SIMS researchers expanded the type of survey questions used to probe topic coverage in order to measure more precisely students' opportunities to learn (OTL). The SIMS experience, in particular, suggests that valid data on instructional content can be obtained from teacher surveys. That research found that mean teacher OTL ratings provided a reasonably good predictor of between-system achievement differences and consequently, had some predictive validity at the level of national education systems (Travers and Westbury, 1989).

However, despite the success of the SIMS strategy in documenting topic coverage, several questions remain about the reliability and validity of content data obtained from national surveys. First, most U.S. surveys ask about topics at a level of generality that either does not differentiate the breadth or depth at which topics cutting across multiple courses are covered (e.g., polynomials, properties of geometric figures) or probes at the level of a single course title (e.g, trigonometry, calculus) and does not give any indication of the specific content of that course. Second, surveys typically do not ask about the amount of time spent on a particular topic--i.e., the number of periods or lessons devoted to the topic. Finally, it is difficult to validate topic coverage in a cost-effective way for indicator purposes. Unless all of a teacher's exams and assignments are collected for an entire school year, these sources cannot provide an accurate picture of the topics covered or the depth of coverage. Textbooks are the obvious alternative because they typically span an entire course and can be collected and coded without burdening teachers. However, given earlier research on elementary mathematics

ERIC

showing that teachers using the same text vary widely in their topic coverage and pacing (Freeman et al., 1983) and the fact that teachers do not typically cover an entire textbook and may supplement it with other materials, textbooks can only be used as a source of validation if information is also available about how they are used by individual teachers.

We tried to address each of these issues in designing our strategy for validating survey data about instructional content. As noted in the previous chapter, our survey contained an expanded list of topics that, in addition to the more general topics included on the NAEP and NELS surveys, included ones at a greater level of specificity. Our survey also asked teachers about the number of periods devoted to each topic. Because of the need to validate topic coverage information that spanned an entire year, we did rely on teachers' textbooks as the primary source for validation. However, we asked them exactly which chapters they covered and how closely they followed the textbook. Only those chapters that teachers indicated they had already covered or planned to cover by the end of the year were coded for content coverage. In addition, although we could not use either teachers' exams (because they covered only one semester) or their assignments and logs (which covered only five weeks) as a primary source for validating topic coverage, we did use them as a secondary source.

Our analysis suggests that there are differences across topics in the accuracy with which their coverage is reported on teacher surveys. Those topics covered in upper-level courses tend to be reported with great accuracy, while the topics reported with less accuracy tend to be those covered in lower-level courses, more general topics, those associated with the mathematics reform movement, and ones that are used as tools in the learning and application of other topics (e.g., graphing, tables and charts). Before presenting the findings from our validation analysis, we provide some examples of the kinds of information that are available from survey data on topic coverage.

DESCRIBING COURSE CONTENT FROM SURVEYS: ILLUSTRATIVE EXAMPLES

Perhaps the most important use for topic coverage data is in describing the distribution of students' opportunities to learn the content associated with a particular course. A number of studies (e.g, McDonnell et al., 1990), including presentations of the

ERIC

SIMS data (Kifer, 1993), have used "box and whiskers" plots to illustrate how topic coverage for a particular course is distributed. We present similar data here, and then elaborate by moving beyond the standard of whether or not a set of topics has been taught as new content to showing how the amount of class time spent on core topics can vary.

Table 3.1 categorizes those topics from the survey that are commonly covered at four different course levels. These four sets of topics are not meant to be exhaustive, but they do represent at least part of the core content for each of the courses listed. Figure 3.1 compares the distribution of the pre-algebra and algebra topics taught as new content in courses below algebra I with that taught in algebra courses. Figure 3.2 makes the same comparisons, but uses as a criterion whether the two sets of topics were taught for six or more periods--i.e., covered in some depth. The line across the middle of each "box" represents the median; the lower and upper boundaries of the box equal the twenty-fifth and seventy-fifth percentiles; the "whiskers" depict the tenth and ninetieth percentiles; and the dots represent outliers beyond the tenth and ninetieth percentiles.

In terms of exposure to core algebra topics, there is little variation in OTL across the algebra I classes in our sample. Even in those classes in the lowest quartile, teachers report that seventy-five percent of the algebra topics are covered and most classes cover 80 percent or more of the core topics. Yet there also seems to be a fair amount of attention in these algebra classes to lower level content, with the typical class also covering 60 percent of the pre-algebra and arithmetic topics as new content. When we examine the distribution of more indepth coverage in figure 3.2, we see considerably greater variation in the proportion of algebra classes in which core topics are covered for longer periods of time. In a typical class, about half of the topics are taught over six or more periods, but a few classes receive almost no indepth coverage of algebra topics while a few at the other end of the distribution spend extended time on most core topics. The boxplot showing the distribution of indepth coverage of pre-algebra and arithmetic topics in the algebra classes indicates that while these classes may be covering a significant proportion of the lower-level topics, they are doing so for relatively brief periods of time. The typical algebra class only covers 20 percent of the pre-algebra

Table 3.1

REPRESENTATIVE TOPICS COVERED AT FOUR COURSE LEVELS

Pre-algebra and Arithmetic

Ratios, proportions, and percents
Conversions among fractions, decimals and percents
Laws of exponents
Square Roots
Applications of measurement formulas (e.g., area, volume)

Algebra I

Polynomials
Linear equations
Slope
Writing equations for lines
Inequalities
Coordinate Geometry
Distance, rate, time problems
Quadratic equations

Algebra II

Polynomials
Quadratic equations
Logarithms
Conic sections
Slope
Sequences
Matrices and matrix operations

Math Analysis/Pre-Calculus

Trigonometry
Polar coordinates
Complex numbers
Vectors
Limits

ERIC

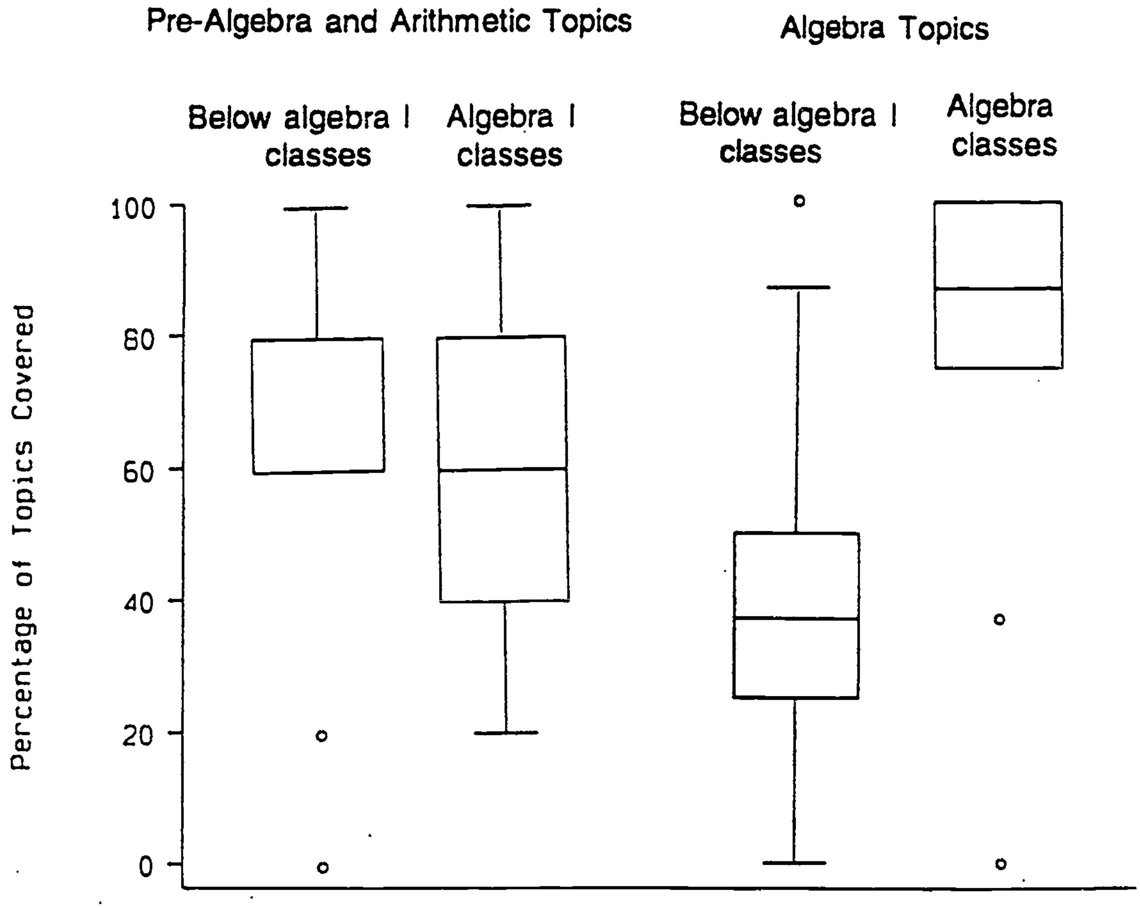

Figure 3.1 -- Proportion of pre-algebra and algebra I topics taught as new content, by course level

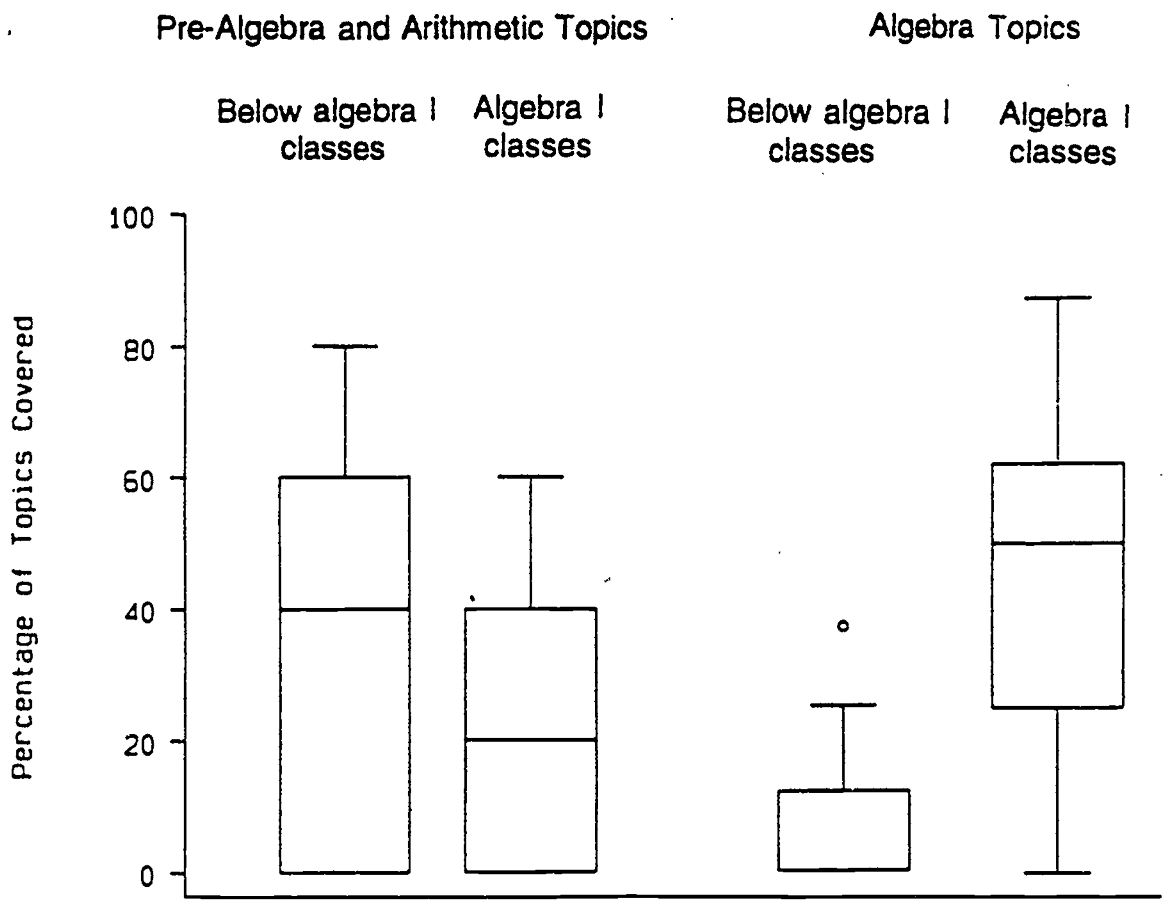

Figure 3.2 -- Proportion of pre-algebra and algebra I topics taught as new content for six or more class periods, by course level

ERIC

topics for more than six periods.

In contrast to the algebra I classes, the pre-algebra classes in our sample show somewhat greater variation in topic coverage, but even at this level, most of the core topics are being covered in the typical class (80 percent in the median class). The proportion of algebra I topics that are included in these classes varies considerably, with those in the top quartile getting some exposure to half of the algebra topics and those in the bottom quartile to about a quarter of the algebra content. However, figure 3.2 illustrates why asking about whether or not a topic was taught without asking about the amount of time spent on it can result in a misleading picture of OTL. As the boxplot showing the coverage of algebra I topics in pre-algebra classes indicates, the median class covers only 12 percent of the algebra topics for six or more periods. Just as the algebra teachers in our sample spent little time teaching lower-level content, the pre-algebra teachers only briefly introduced their students to algebra topics.

Figures 3.3 and 3.4 present the same kind of information for algebra II and math analysis classes. What is striking about these classes is that in each case there is little variation in coverage of the core course content. In the typical algebra II class, about 80 percent of the algebra II topics are covered and about the same proportion of math analysis topics are covered in the math analysis courses. Similarly, in both courses, most of the core topics are covered for six or more periods. Still there is considerable overlap in topic coverage between the two courses, with the median algebra II class covering 40 percent of the math analysis topics (20 percent for six or more periods), and the median math analysis class covering 71 percent of the algebra II topics (35 percent for six or more periods).

This presentation of topic coverage expands on past uses of course content data to illustrate how knowing the amount of class time spent on a set of topics provides a more accurate measure of students' opportunity to learn. With this additional information, we found that some students are receiving indepth instruction on core topics, while others are only briefly introduced to them. Comparing topic coverage across course levels also allows us to estimate the distribution of course-level content that students are receiving, as compared with their exposure to topics that are either above or below the level of the

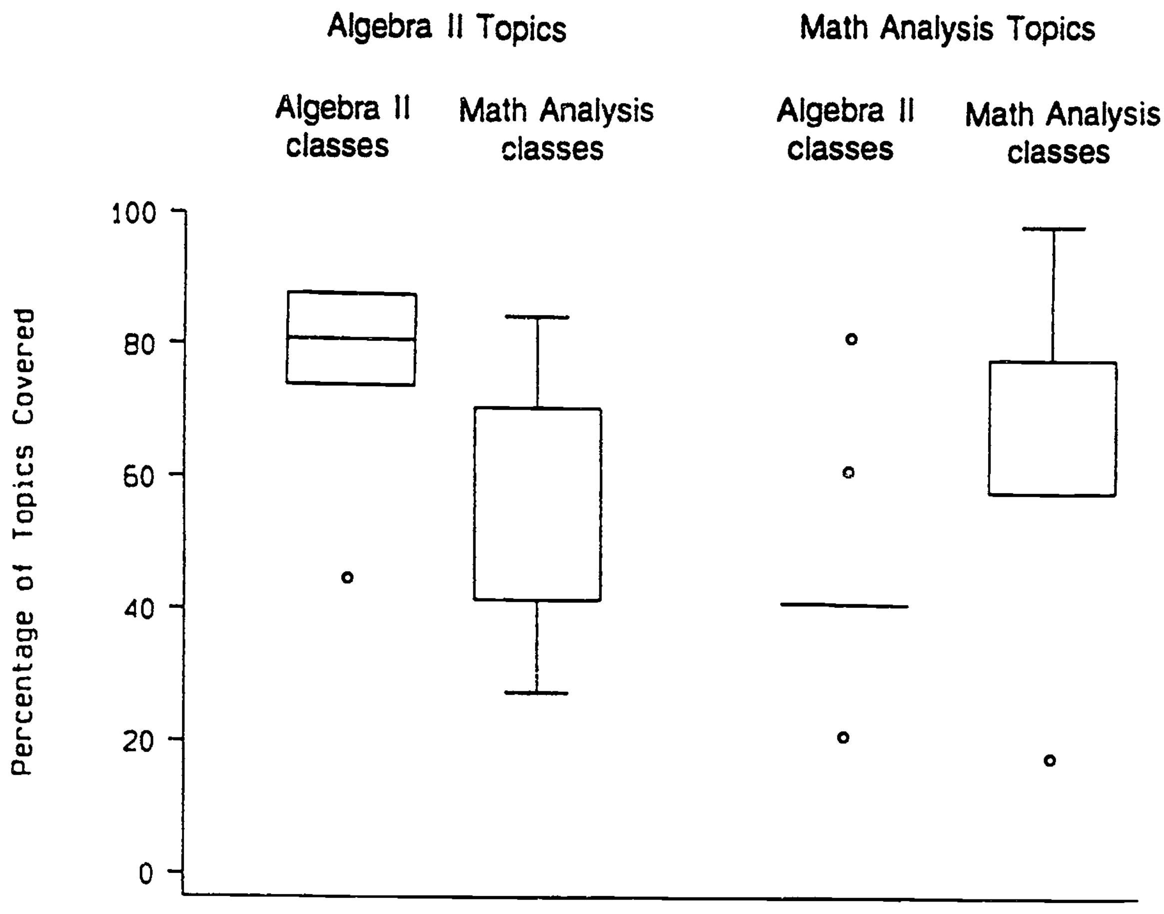

Figure 3.3 -- Proportion of algebra II and math analysis topics taught as new content, by course level

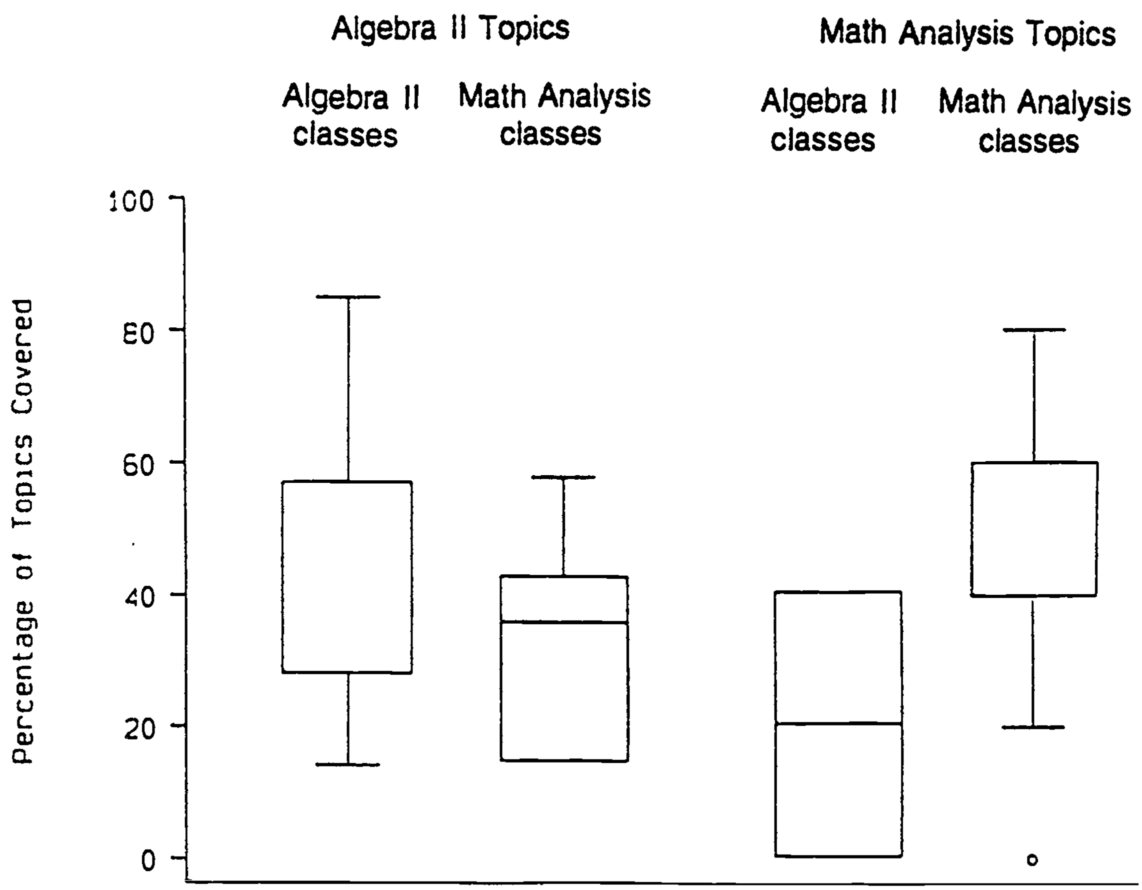

Figure 3.4 -- Proportion of algebra II and math analysis topics taught as new content for six or more class periods, by course level

ERIC

course. Using indicator data on topic coverage in these ways, can provide a quite thorough depiction of instructional content. However, the quality of that picture depends on the accuracy of the survey data from which it is drawn.

CONSISTENCY BETWEEN THE SURVEYS AND THE COURSE TEXTBOOKS

For those chapters of their textbooks that teachers indicated they had covered or planned to cover before the end of the year, coders counted the number of lessons devoted to each of the topics listed on the post-data collection survey.[10] These counts were then converted into the six response categories in question 11 of the survey.[11] using the algorithm outlined in the footnote to Table 3.2.

Across all topics, the average rate of direct agreement between the surveys and textbooks is 42 percent and within one survey response category, 72 percent. The average level of agreement suggests that although survey data may not provide a very precise picture of the time spent on different topics, such information is reasonably accurate at the level of being able to ascertain that a topic has been taught not at all or for only a few periods, for a week or two, or for several weeks. Given that most OTL measures tend to be fairly crude (i.e., typically reporting whether or not a general topic has been covered with no information about the time spent on it), being able to report topic coverage at this level of specificity represents a significant improvement.

But the mean rate of agreement masks significant differences across topics. Table 3.2 lists those topics (15 of the 40 included on the two forms of the survey) for which the rates of agreement were below the mean. Table 3.3 lists the eight topics for which the rates of agreement were the highest. Five of these topics[12] are covered primarily in

---

[10]The instructions to the coders defined a textbook lesson as a subsection of a chapter that typically consists of a two to three page "spread," with a page or two of explanation followed by a set of exercises. Coders were instructed to include all substantive sub-sections and to omit enrichment, computer, review and chapter test sections.

[11]The six response categories were *0 periods, 1-2 periods, 3-5 periods, 6-10 periods, 11-20 periods*, and *>20 periods*.

[12]The five topics covered in upper-level courses are: *calculus, measures of dispersion, integration, discrete math*, and *vectors*.

Table 3.2

Consistency Between Topic Coverage As Reported on Surveys and As Coded from Textbooks*

Topics Where Direct Agreement Is <42% and Agreement Within One Survey Response Category Is <72%			
	% Direct Agreement	% Within One Category	Possible Explanation For Inconsistency
Linear Equations	9.1	42.4	Low inter-coder reliability (K=.259)
Conversion among Fractions, Decimals, Percents	20.6	70.6	Use vs. specific focus of teaching
Conic Sections	24	60	
Polynomials	21.7	55	Low inter-coder reliability (K=.359)
Graphing	25	66.7	Use vs. specific focus of teaching
Inequalities	27.3	69.7	Low inter-coder reliability (K=.324)
Tables & Charts	25	66.7	Use vs. specific focus of teaching
Proportional Reasoning	25.9	65.5	--Lack of understanding or common meaning among respondents --Low inter-coder reliability (K=.067)
Patterns & Functions	22.9	64.6	Lack of understanding or common meaning among respondents
Ratio, Proportion, Percents	25.7	54.3	Low inter-coder reliability (K=.189)
Sequences	29.6	59.3	
Slope	26.7	68.3	Low inter-coder reliability (K=.324)
Math modeling	28.3	50.9	Lack of understanding or common meaning among respondents
Estimation	33.3	70.6	Use vs. specific focus of teaching
Matrices	42.3	69.2	

* Textbooks were coded for the number of lessons in which a topic was covered. Because textbooks divide material differently and include varying numbers of lessons, the number of textbook lessons covered by teachers ranged from 34 to 181, with a mean of 93.6 and a standard deviation of 29.8. In order to standardize across texts and to make valid comparisons with teachers' reports about the number of periods spent on a topic, the number of lessons on a given topic was divided by the total number of lessons and multiplied by 140, which is an approximation of the total number of periods of instruction in a given academic year. The resulting number was then converted into one of the six response options for reporting topic coverage on the survey (*1=0 periods, 2=1-2 periods, 3=3-5 periods, 4=6-10 periods, 5=11-20 periods, 6=>20 periods*).

Table 3.3

Consistency Between Topic Coverage As Reported on Surveys and As Coded from Textbooks*

Topics Where Direct Agreement Is >50% and Agreement Within One Survey Response Category Is >80%		
	% Direct Agreement	% Within One Category
Calculus	85.2	88.9
Measures of Disperson	80.8	92.3
Integration	80.8	96.2
Discrete Math	76.9	92.3
Growth & Decay	66.7	86.7
Vectors	63	88.9
Probability	59	80.3
Statistics	50.8	82

* The process by which the textbook data were recoded to be comparable with the survey data on topic coverage is described in the footnote to table 3.2.

ERIC

upper-level courses, so only a minority of the teachers in the sample reported spending any time on them. In addition, because the scope and sequence of upper-level courses tends to be more precisely defined (e.g., because of the requirements of Advanced Placement tests and the narrower focus of the topics covered), teachers may be able to estimate more precisely the amount of time spent on a topic.

We identified six possible reasons for the low rates of agreement. The first was the algorithm we used in converting the continuous data on topic coverage, as reflected in the textbooks, into the same categories that teachers used on the surveys. We knew that textbook lessons as a unit of analysis may not be exactly comparable to periods of instruction, and that the number of days available for instruction varies by school, and even for classes within the same school (e.g, depending on when assemblies, standardized testing and the like are scheduled). Consequently, we tried standardizing the two measures of topic coverage in a variety of different ways (e.g., by raising or lowering the approximate number of periods of instruction over the year, by examining only textbook lessons and topics reported on the survey as having already been covered), but none of these transformations produced significantly different rates of agreement.

Second, although the double-coding of textbooks indicated that coder error had been kept to within an acceptable level,[13] some topics were more prone to coder disagreement than others. Six of the 15 topics in Table 3.2 had rates of interrater agreement that are only slight or fair (i.e., a kappa statistic of $<.40$)[14], and these were the lowest among the 40 topics coded. Although low inter-coder reliability was only found to be a problem for a few topics, it does suggest that coders may need more training and ongoing monitoring than we provided.

---

[13]As noted in chapter 2, on average, two coders agreed on the exact number of lessons devoted to a topic 58 percent of the time, were within one lesson 75 percent of the time, and within two lessons, 85 percent of the time.

[14]A kappa statistic is a measure of interrater agreement when there are two unique raters and two or more ratings It is scaled to be 0 when the amount of agreement is what would be expected to be observed by chance and 1 when there is perfect agreement. Moderate levels of agreement are conventionally interpreted as .41-.60, substantial as .61-.80, and almost perfect as .81-1.00. All but nine of the topics coded from the textbooks had moderate levels of agreement or higher, with 19 of the topics at the substantial level or above.

A third explanation emerged from our follow-up interviews with a sub-sample of the teachers. We returned to talk with them because we noted internal inconsistencies in their surveys (most notably between their reports about instructional goals and activities) that suggested there was a lack of common understanding for some terms associated with the mathematics reform movement. Consequently, we asked teachers how they defined the topics we found to be the most problematic. Three appeared to present special problems for respondents, and one of these--*proportional reasoning*--was also the one with the lowest level of interrater agreement. Teachers told us that the term proportional reasoning was vague; some reported that they had never seen the two words combined. For *math modeling*, several teachers participating in the group interviews volunteered that they had no idea what the term meant, even though it appears in key mathematics reform documents (National Research Council, 1989; California Department of Education, 1992). For another topic, *patterns and functions*, interviewees argued that the two concepts should be separated because they are not parallel or necessarily linked concepts. In this instance, the newer reform literature (California Department of Education, 1992) seems to agree with our teacher respondents, and argues against NCTM's (1989) joining the two concepts. The reason is that patterns play a broadly applicable role in many or perhaps all strands of mathematics, while functions comprise one specific way of generalizing an observed pattern. Although the problem of teachers either not understanding the meaning of a term or interpreting it differently across respondents is considerably greater for instructional activities and goals than for topics, these examples do suggest that some survey data cannot be validly interpreted during a time in which language and accompanying practice are in transition. While only a few topics may fall in this category, they are the ones of potentially greatest interest for charting trends in curricular reform.

Fourth, for four of the topics with low rates of agreement, we found that teachers reported spending greater amounts of time on teaching them than the coders estimated. These four topics--*conversion among fractions, decimals, and percents; graphing; tables and*

ERIC

charts; and *estimation*--have in common that they are tools or building blocks that students can draw upon in working problems on other substantive topics. For example, some geometry textbooks have students record their measurements of geometric figures in a table format. Although *tables and charts* is not the specific focus of teaching during such exercises, it is being used by students. We believe that teachers' overestimation of their coverage of these topics stems from their not making a distinction between having students use the concept while working on other topics and having the topic as the primary focus of a substantive lesson. This particular problem can be addressed by including clearer instructions in survey prompts.[15]

In designing this study, we assumed that there are two other factors which affect the validity of topic coverage data obtained from teacher surveys. One is the high level of generality that characterizes the topics included on most national surveys. Because we assumed that specific topics will yield more valid data, as well as a more detailed picture of students' opportunity-to-learn, we included a greater number of specific topics on our post-data collection survey. Several of those topics were elaborations of the NELS topics. As Table 3.4 indicates, more specific topics had higher rates of agreement between the surveys and textbooks than did the general NELS topics. We recognize that there are trade-offs associated with including longer topic lists, particularly on surveys that must serve multiple purposes in addition to gathering data on curriculum. But our research suggests that, despite the potentially greater teacher burden, national surveys will need to be more comprehensive if they are to provide valid data on topic coverage.

A final factor likely to affect the reliability and validity of survey data on topic coverage is the time frame over which teachers are asked both to recall what they have already taught and to estimate what they will teach over the rest of the year. We assumed that such data are likely to be considerably less precise than if teachers are

[15]As indicated in Table 3.2, we found that these four reasons helped explain most of the lowest rates of agreement between the textbooks and the surveys. However, for three topics--*conic sections, sequences*, and *matrices*--none of these reasons apply. The rate of interrater agreement was at the moderate level or above for these topics; teacher respondents agreed on their meaning; and they did not systematically overestimate their coverage of these topics as they did for those that might be classified as tools.

asked to report on topic coverage concurrently with their actual teaching. We were able to test that assumption by comparing teachers' reports of topic coverage on their daily logs with the assignments they made over the same five weeks. Eleven topics were reported as being covered over this period by at least a third of the teachers. Across those topics, the rate of direct agreement was 58 percent and 83 percent within one category, using the same response categories as the survey. To ensure that this higher rate of agreement was not an artifact of converting the continuous data from the logs and assignments into the survey categories, we also calculated the rate of agreement between the exact number of periods teachers reported covering a topic and the number of times coders identified it as being included on the assignments. We found that the average rate of agreement within one class period or assignment was 40 percent and 59 percent within two. The extent of improvement in the quality of data collected simultaneously with the teaching of a topic is illustrated by the fact that eight of the 11 topics are also ones that had the lowest rates of agreement between the survey and the text. Using the log and assignment data multiplied the rate of agreement by between one-and-a-half and two-and-a-half times for these topics.

Clearly, asking teachers to report on their topic coverage concurrently with their teaching of the content is not a feasible strategy for routine data collection in national surveys. The gains in improved reliability would be offset by an incomplete picture of the content presented to students throughout the year. In addition, having a large number of teachers report on a daily basis for even several weeks might increase the costs of national surveys and would most likely result in a lower response rate overall.

CONCLUSIONS

Our analysis suggests that teacher survey data can provide a reasonably accurate of topic coverage. If the standard is knowing whether or not a topic has been taught and if it has been taught, whether it has been covered over several periods, for a week or two, or for several weeks, then teacher self-reports are reliable. However, our data provide a strong rationale for including more specific curricular topics on surveys. Not only do they provide a more detailed and comprehensive picture of students' opportunity-to-learn, but teachers' reports on these topics are more reliable than their reports about general topics which encompass multiple sub-topics and make precise time estimates difficult. We recognize the trade-offs in requesting more detailed information about topic coverage, but we would argue that the topics currently included on national surveys such as NELS provide data that are too general to be useful, particularly in measuring OTL. Consequently, the gains in both reliability and validity may more than offset the additional burden.

In addition to the need for more detailed, enhanced topic lists on national surveys, our artifact analysis suggests that validation studies are necessary to pinpoint the sources of measurement problems. One area that will continue to be problematic is the lack of common agreement on the meaning of key terms associated with the mathematics reform movement. Such terms need to be included on national surveys to chart trends in topic coverage, but without accompanying validation studies, the data are likely to be misinterpreted. Consequently, the use of indepth interviews and focus groups to supplement artifact analyses will help in identifying the different understandings that teachers hold of concepts central to expected changes in mathematics teaching. But even independent of the current flux in curricular practices, validation studies are necessary. By collecting detailed data from multiple sources over shorter periods of time (e.g., through daily logs and assignments), such studies can provide a benchmark against which to judge the reliability of routine survey data that require teachers to recall and estimate topic coverage over longer periods of time.

Chapter 4

INSTRUCTIONAL STRATEGY

Instructional strategy is a multi-faceted dimension of curriculum that is considerably more difficult to measure than instructional content. Part of the reason lies in its scope. Embodied in this concept are all the various approaches used in the teaching and learning process. It includes what teachers do (e.g., lecture, lead discussions, work with small groups) and what students do (e.g., work individually or in groups, work on projects, use manipulatives). But it also includes the type of work that students are assigned--the focus and format of that work, how it is evaluated, and the level of understanding expected of students.

Instructional strategy is also difficult to measure because surveys typically cannot capture the subtle differences in how teachers define and use different techniques. For example, one teacher might lecture directly from the textbook, and do most of the talking in the class. Another might draw on material from sources other than the text and engage students in lively give-and-take exchanges. Without a detailed set of survey probes, however, both teachers are likely to report that they spend most of the period lecturing. Even a detailed survey would likely fall short in representing these two classrooms because it could not adequately measure the nature of the interaction between students and the teacher.[16] Yet someone observing those classrooms would identify two very different kinds of instruction.

Despite these major limitations, however, there is still much that survey data can tell us about instructional strategy. Such data can describe the major dimensions of classroom processes and how they vary across course-levels and types of schools. National survey data, collected on a periodic basis, can document trends in teachers' use of generic instructional strategies. Such information is important in determining whether or not teaching is changing in ways consistent with the expectations of curriculum reformers and their policymaker allies.

---

[16]This shortcoming of survey instrumentation also applies to artifact data. In fact, artifact data are particularly weak in their ability to portray instructional strategies that do not involve written work of some type.

PORTRAYING INSTRUCTION FROM SURVEYS: ILLUSTRATIVE EXAMPLES

The clearest picture of instruction that emerges from our survey data is teachers' reliance on a few strategies that they use frequently. A large proportion of teachers reported engaging in traditional activities such as lecturing (87 percent) and correcting or reviewing homework (86 percent) on a daily basis, while the majority reported engaging in activities consistent with the mathematics reform movement on only an infrequent basis or not at all--e.g., 65 percent of teachers reported having student-led discussions once or twice a semester or not at all; 61 percent rarely or never discussed career opportunities in mathematics, and slightly fewer than half of our respondents (49 percent) have their students work in small groups at least once or twice a week.

Calculator use in these classrooms is very high, with 74 percent of the teachers reporting that students use them almost daily. Usage is just the opposite for computers, however. Students use computers on a daily basis in less than 2 percent of the courses, and in over half of the classes (52 percent), computers are never used to solve exercises or problems.[17] Most teachers reported that the majority of class time is spent in direct instruction, with student discipline and administrative tasks such as taking attendance consuming less than 10 percent of their in-class time.

The tendency in national indicator reports such as those produced by NAEP (e.g., Mullis, et al., 1991; 1994) has been to focus on single questionnaire items, examining each teaching strategy separately rather than seeking to understand how teachers link discrete strategies to create instructional repertoires. Given that teachers rarely use just one strategy and typically rely on several even in the same lesson, reporting on an item-by-item basis fails to produce a coherent picture of instruction. Consequently, we probed our survey data to see if we could identify different instructional repertoires in which teachers combine a number of separate strategies. Our first approach consisted of grouping instructional techniques according to the strategies advocated in reform

[17]The low incidence of computer use reported by the mathematics teachers in our sample is similar to the level of reported use found by Weiss (1994) in her national survey of science and mathematics teachers. Fifty-six percent of the high school mathematics classes in her sample never use computers. The level of teacher lecture, textbook usage, and small group work in our sample is also similar to the patterns documented by Weiss.

documents such as the NCTM Professional Standards (1991) and the California mathematics framework (1992). We also created a list of techniques which seemed to represent the more traditional teaching repertoire to which the reform documents were reacting. Table 4.1 shows the two groups of strategies. To test the consistency of these groupings, we scaled them, and found that they cohered reasonably well with an alpha of .72 for the reform scale and .62 for the traditional scale. The parts of the scale that were intended to operate in a negative direction (e.g., not lecturing as part of the reform repertoire) did in fact reverse as expected.

We also conducted a factor analysis as another way of identifying instructional strategies that occur together. Three factors emerged that seem to have substantive meaning. The first, shown in Table 4.2, is dominated by discussion strategies with a strong emphasis on the role of students in class discussions. The second has a clear demonstration component. Student participation strategies are part of this repertoire, but it is much more teacher-directed than the first one. The third factor, with only two components, is the closest to what would be considered traditional teaching with the teacher lecturing and students responding to the teacher's questions.

The lack of variation in classroom practice across the teachers in our sample is the primary reason why the results of the factor analysis indicated that most of the instructional strategies included on our survey do not fit into a common factor space. Nevertheless, the coherence of the reform and the traditional practice scales and the high face validity of the factors that did emerge suggest that future efforts to link instructional strategies and student outcomes should move away from separate analyses of single questionnaire items and focus greater attention on identifying and understanding different instructional repertoires.

The picture of instruction that emerges from our survey data is quite consistent across course levels. As might be expected, those teaching algebra I and courses below that level had students practice or drill on computational methods more frequently than teachers in higher level courses. In addition, the number of minutes per day of homework that teachers assigned was significantly greater for higher level courses, with the mean ranging from only 19 minutes per day in courses below algebra I, to about a

ERIC

Table 4.1
Instructional Repertoire Scales

	Reform	Traditional
Lecture	−	+
Students respond orally to questions		+
Student-led discussions	+	
Teacher-led discussions		+
Review homework		+
Students work individually		
Students give oral reports	+	
Administer a test	−	
Administer a quiz		
Discuss career opportunities		
Small groups work on problems	+	
Whole class discusses small groups' solutions	+	
Students read textbooks		+
Teacher summarizes lesson's main points		+
Students work on next day's homework		+
Students work on projects in class	+	
Teacher demonstrates an exercise at board	−	+
Students work exercises at board		
Teacher uses manipulatives to demonstrate a concept		
Students work with manipulatives	+	
Students practice computational skills		+
Students work on problems with no obvious method of solution	+	−
Students use tables and graphs	+	−
Students use calculators		−
Students use computers	+	
Students respond to questions that require writing at least a paragraph	+	−
	$\alpha = .72$	.62

Table 4.2
Instructional Repertoire Factor Matrix*

	Factor 1	Factor 2	Factor 3
Student-led discussions	.51	-.18	.04
Teacher-led discussions	.67	-.02	.34
Small groups work on problems	.83	-.02	-.18
Whole class discusses small groups' solutions	.74	.07	.14
Students use calculators	.54	-.19	-.18
Students respond orally to questions	.16	.78	.89
Administer a test	.11	-.85	-.06
Teacher demonstrates an exercise at board	-.30	.54	.27
Teacher uses manipulatives to demonstrate a concept	.41	.51	-.22
Students work with manipulatives	.27	.55	-.04
Lecture	-.15	.01	.80
Review homework	-.28	-.15	.20
Students work individually	-.16	.07	-.07
Students give oral reports	.21	.39	-.07
Administer a quiz	.14	-.01	.30
Discuss career opportunities	.23	-.00	.39
Students read textbooks	.00	.10	.24
Teacher summarizes lesson's main points	.02	.46	.40
Students work on next day's homework	-.16	-.11	.24
Students work on projects in class	.12	.43	-.16
Students work on exercises at board	.20	.25	.09
Students practice computation skills	-.11	-.22	-.08
Students work on problems with no obvious method of solution	.44	.36	.11
Students use tables and graphs	.43	.19	-.18
Students use computers	.06	.07	-.22
Students respond to questions that require writing at least a paragraph	.38	.31	-.25
Eigenvalue	4.62	2.93	2.14

* Although six factors were extracted, the three primary factors are presented here because the scree diagram indicated that the remaining factors did not explain a substantial, additional proportion of the variance, and they were not readily interpretable on substantive grounds.

half hour a day for algebra I (32 minutes), to almost an hour a day for calculus (55 minutes). Beyond these differences, however, there were few other significant differences across course levels. Teachers in higher level courses were just as likely as those in lower level courses to lecture frequently, have students work on their next day's homework in class, and then correct or review that homework. Similarly, the infrequency of strategies such as student-led discussions and small group work was quite similar across course levels.

One course, however, does seem to rely on different teaching strategies. Although the number of calculus classes in our sample was too small to make any generalizations, those classes did differ from the other courses in the sample in several major ways. Calculus teachers reported lecturing less frequently and relying more on small group work by students. But with this exception, the similarities across courses in our sample are far more striking than the differences.

CONSISTENCY BETWEEN THE SURVEY AND THE ARTIFACTS

Since instructional strategy is the dimension of curriculum least amenable to validation through written artifacts, we were limited in our ability to measure the consistency of survey responses with other data. However, for 14 of the 26 instructional practice items listed on the survey, we could compare teachers' survey responses at the end of the semester with their daily log entries during the five weeks of artifact data collection. We were also able to compare teachers' survey responses about the format and other characteristics of their exams and quizzes with their artifacts. Finally, we were able to compare teachers' actual homework assignments with their responses to a survey question about the characteristics of those assignments.

Logs and Surveys

Table 4.3 shows the rate of exact agreement between the surveys and the logs on the frequency with which teachers reported engaging in a variety of instructional activities. The level of agreement within one survey response category is also

ERIC

Table 4.3

Consistency Between the Reported Frequency of Instructional Activities on the Logs and the Surveys

	% Direct Agreement	% Within One Survey Response Category
Lecture	57.4	96.7
Have students respond orally to questions	46.7	90
Have teacher-led whole group discussions	35	86.7
Correct or review homework in class	38.3	85
Have students work individually on written assignments	43.3	83.3
Administer a test	60	83.3
Have students work with other students in small groups	50	100
Have students work on next day's homework in class	34.6	61.5
Demonstrate exercise at the board	50	84.6
Have students work exercises at the board	59	91.8
Use manipulatives to demonstrate a concept	52.5	83.6
Have students work with manipulatives	48.3	86.7
Have students use a calculator	47.2	90.6
Have students work on a computer	49.1	86.8
Mean rate of agreement	48.0	86.5

ERIC

reported.[18] Given that the data in this table compare information that teachers provided at three times over the course of the semester (the last time within one week of completing the survey) with their responses to the survey completed at the end of the semester, the rate of direct agreement is quite low.[19]

The relatively high rate of agreement within one survey response category does suggest, however, that the problem may lie in how the survey response categories were constructed. The distinctions among them may not have been sufficiently discrete or meaningful to respondents. To check this possible explanation, we compared the mean frequency of instructional activities, as reported on the logs, with the category teachers used in responding to the survey. We found that the response options teachers used on the surveys did not always reflect actual differences in the frequencies of instructional activities as they had reported them on the logs. The most common problem was the lack of significant differences in the frequency of activities as reported on the logs for the survey response categories, *almost every day* and *once or twice a week*, and for *once or twice a week* and *once or twice a month*. For eight of the 14 activities for which a comparison was made, the mean for two of the survey categories was virtually the

[18]In order to compare the survey responses with the log entries, we had to convert the continuous data from the logs--i.e., how many times over the 25 days of data collection teachers reported engaging in an activity--into the categorical response options used in question #13 on the survey. Anything that was done 60 percent or more of the time (> 15 days) was recoded as *almost every day*, any activity that occurred 25-59 percent of the time (5-14 days) was recoded as *once or twice a week*, and 24 percent or less (< 5 times) was recoded as *once or twice a month*. We included *never* as a comparison category, but did not have a comparison category for *once or twice a semester* for the log data.

[19]In comparing the survey and log data, we chose to use the rate of direct agreement and the rate of agreement within one survey response category as our measure of consistency. Another recent study (Porter et al., 1993) that made similar comparisons between log and survey data used correlation coefficients as the measure of consistency. That study found the correlations between log and questionnaire data to be "substantial," (2-31) and concluded that "the validation results were very encouraging" (A-5).

Although the correlations for seven of the 14 instructional practice items on our survey and log were greater than .50 (and significant at the .01 level) and only two items had correlations below .30 (with one significant at the .05 level and the other nonsignificant), we did not find this analysis to very informative. Correlation coefficients conflate matches and mismatches across categories in a way that makes it difficult to retrieve information about specific patterns of responses to the two types of data collection instruments. A clearer and more intuitively attractive way to compare the two is to examine how close the agreement is between the two indicators. The percentage agreement statistic measures that level of consistency directly.

ERIC

same.[20] Although respondents reported no problems in using these response categories, the log data suggest that teachers who had engaged in an activity with the same frequency used different categories in reporting on it on their surveys. The response categories that are the most problematic are *almost every day* and *once or twice a week*, with the log data suggesting that reliable distinctions cannot be made between these two categories, based on survey data.

Exams and Surveys

Another aspect of instructional strategy that we were able to validate from the artifacts was the format and characteristics of teachers' exams. Teachers were asked to indicate the proportion of their tests and quizzes that were *multiple choice*, *short-answer*, *essay*, and *open-ended problems*. In addition, they were asked to indicate the proportion of their exams that included items with certain characteristics such as requiring students to describe how to solve problems or problems with more than one possible answer or one possible approach. Figures 4.1, 4.2, and 4.3 compare the means for the survey responses and the artifact coding.

On six of the 15 questions that teachers were asked about their exams, the level of agreement between teachers and coders was 90 percent or more. There was high agreement between teachers and coders in the proportion of exam items that were multiple choice (with the difference in means between the two sources less than 3 percent) and those that were essay (the difference in means was 3 percent). Similarly, there was high agreement on the proportion of test items requiring the use of tabular or graphical data, the proportion requiring students to describe how to solve problems, and the proportion of problems with more than one answer. However, there were major disagreements between survey respondents and coders about the proportion of exam items that were short-answer (a difference between the two sources of about 50 percent) and that were open-ended problems (a 51 percent difference).

[20]For example, for *lecture*, the mean over the twenty-five days of data collection for teachers indicating *almost every day* on their survey was 14.8 and for those reporting *once or twice a week*, 12.6. For *teacher-led discussion*, the means were 12.8 and 12.1 respectively; *demonstrate an exercise at the board*, 15.4 and 14. Similarly, for *administer a test*, the mean for *once or twice a week* was 4.2 and for *once or twice a month*, 4.3.

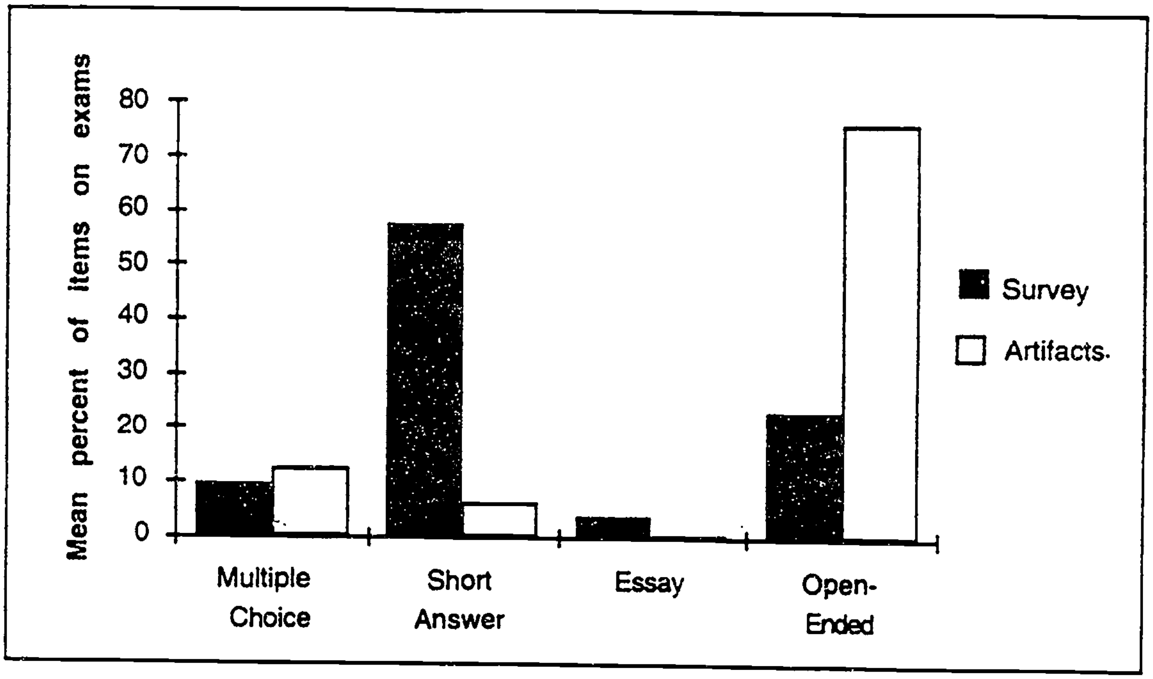

Figure 4.1 - Comparison of Item Formats on Exams

BEST COPY AVAILABLE

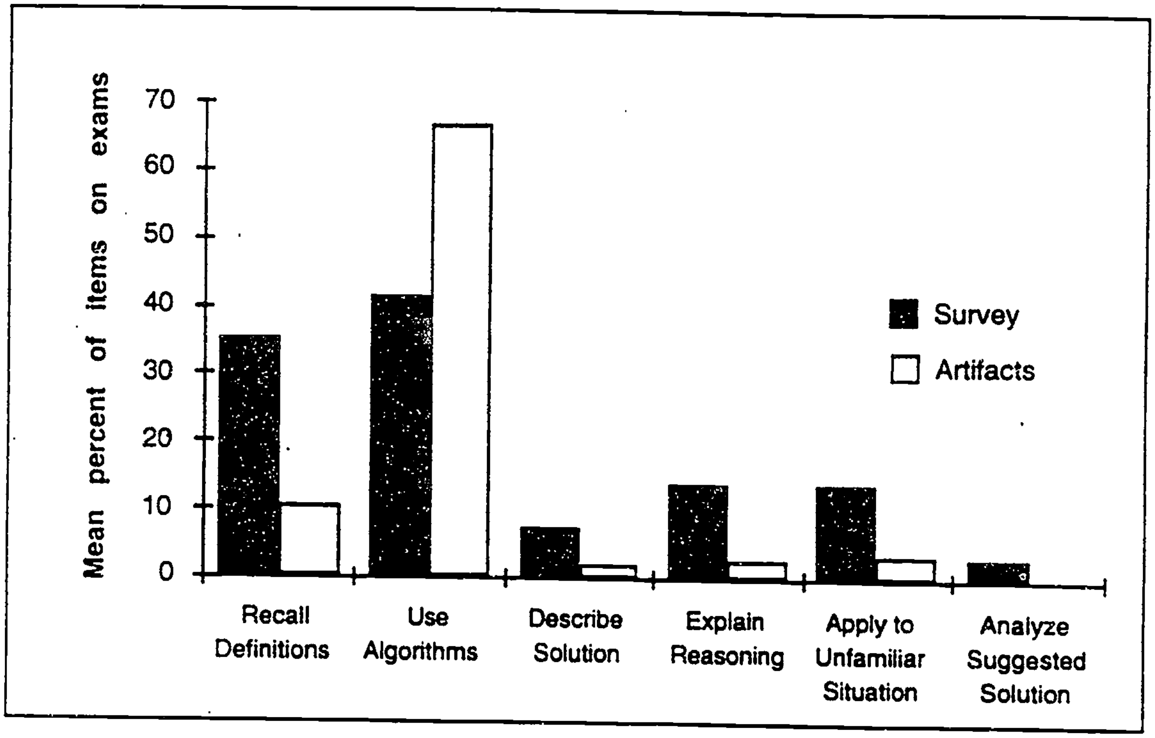

Figure 4.2 - Comparison of Exam Item Characteristics

ERIC

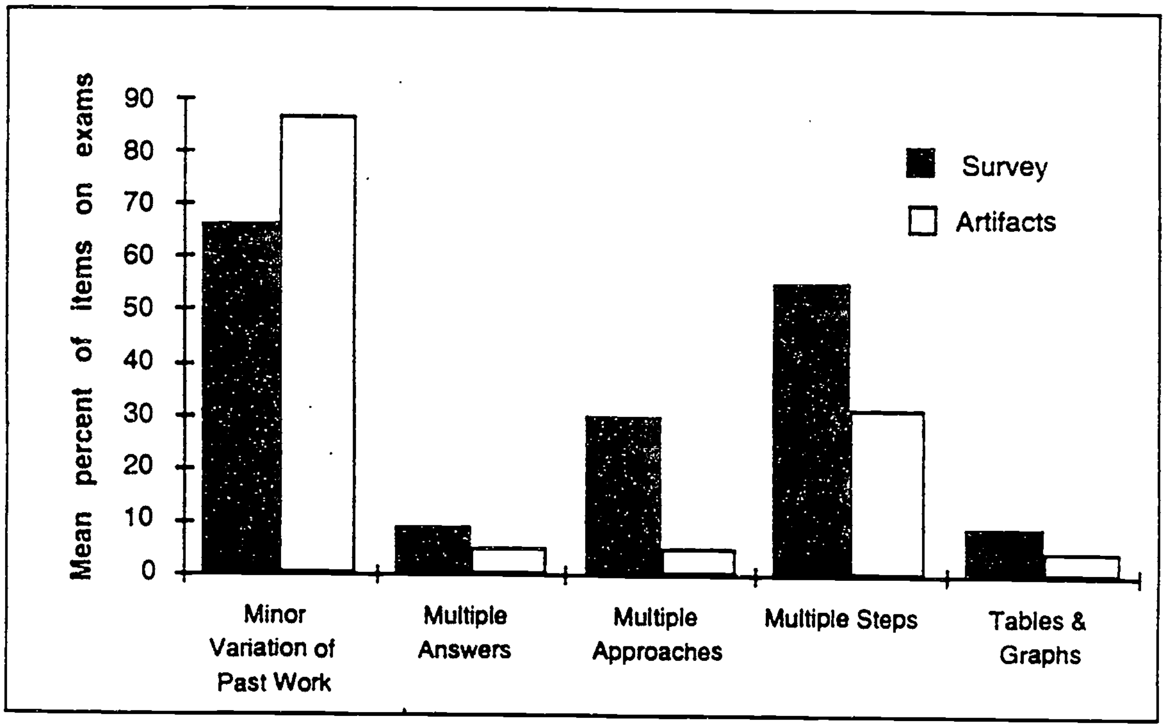

Figure 4.3 - Comparison of Exam Problem Types

ERIC

Other questions where the level of disagreement between the surveys and the artifacts was greater than 20 percent were the proportion of exam items: requiring students to recognize or recall definitions; requiring the use of algorithms to solve problems; that are minor variations of homework, class exercises, or problems; with more than one possible approach; and requiring more than one step to reach a solution. As illustrated in the histograms, the coding of teachers' exams presents a more traditional picture of their approach to evaluating students than they reported on their surveys. According to the artifacts, teachers were less likely to include items that required students to describe how they solved problems, explained their reasoning, or applied concepts to different or unfamiliar situations than they indicated on their surveys. Similarly, the exams evidenced a smaller proportion of items with more than one possible answer, more than one possible approach, or that required more than one step to reach a solution than respondents estimated on their surveys.

The survey response categories related to exam characteristics that were particularly problematic were short-answer, open-ended problems, and the difference between the two. Even though the survey instrument defined open-ended problems as those *where students generate their own solutions*, teacher respondents and coders viewed the exam formats quite differently, with teachers tending to classify as short-answer those items that coders categorized as open-ended problems. In classifying exam items, coders used a narrow definition for short answer--viz., a question requiring students to complete a sentence or fill-in-the blanks. However, our follow-up interviews indicated that while some teachers had interpreted the term more broadly, they also differed in their definitions. For example:

> Short-answer is giving students a specific question to answer. On the other [open-ended problems], students can go in different directions and they are graded on how indepth their answer is. (Math B teacher)

> A short-answer is when students are following standard procedures. I use open-ended when I'm introducing new topics and I don't give students a way to do it. I give them a good background on what they should be finding, but I don't guide them. An open-ended problem ....is more about what students are expected to do, than the format of the test. Just because a test doesn't have a blank to fill in for the answer doesn't make it open-ended. (Calculus teacher)

ERIC

What's the difference among short-answer, essay, and open-ended? They're all the same. (Intermediate math teacher)

The differing interpretations of these seemingly straightforward terms, as evidenced in the discrepancies between the survey responses and the exam coding and in the follow-up interviews, illustrate the need to define a number of survey items more precisely. In the case of questions about exam format, a more precise set of response options might be:

- Multiple-choice
- Problems where students generate a solution and show their work, but no written explanation is required
- Problems where students generate a solution, show their work, and are also expected to explain their work in writing

This set eliminates the ambiguity inherent in distinguishing among short-answer, essay, and open-ended items, while making a clear distinction between multiple choice and constructed responses and within the constructed response category, between answers that require written explanation in addition to mathematical calculations.

Homework Assignments and Surveys

Teachers' daily homework assignments are a final source of validation about their instructional strategies. Teachers were asked on the survey how often they assigned certain types of homework, and their actual homework assignments were then coded to determine the extent to which they reflected these characteristics. For those assignments that were not done entirely during class time, the rate of direct agreement between the survey responses and the artifact coding was 48 percent and within one response category, 73 percent.[21] Behind this overall level of agreement is the same pattern that

[21]Of the 1407 individual assignments in our artifact sample, 230 (16 percent) were worked on only during class, 389 (28 percent) were worked on by students both during and outside class, 223 (16 percent) were done only outside class, and for the remaining 563 (40 percent), teachers did not designate where the assignments were done. In comparing artifacts with the responses to question 21, we chose to include all assignments, except those that were done only during class time, because we assumed most of the undesignated ones were homework assignments. However, we also checked the rate of agreement between the survey and those assignments that were done either completely outside class or worked on both during and outside of class. The rate of direct agreement was 42 percent and within one response category, 69 percent--essentially the same pattern as for the larger set of assignments.

was evident for other types of instructional strategies. The teachers in our sample rely on only a few types of assignments, and while they report the predominance of these in their survey responses, teachers still indicate greater variety in their assignments than were identified in the artifacts. Ninety-one percent of the assignments in our artifact file are either exercises or problems from the textbook (72 percent) or exercises or problems from worksheets (19 percent). Similarly, 83 percent of the teachers report on their surveys that they assign textbook problems at least once or twice a week, and 72 percent report assigning worksheet problems with the same frequency. A substantial proportion of teachers also report that they *never* give homework assignments that require students to write definitions of concepts (40 percent), solve problems for which there is no obvious method of solution (23 percent), or extend results established in class (29 percent). Nevertheless, another sizeable group of teachers reported using these more innovative homework strategies and going beyond just textbook problems--e.g., over half report assigning homework problems with no obvious method of solution at least once or twice a month. However, the artifacts present a picture of homework assignments that are more traditional with considerably less variety in the type of tasks required of students. The artifacts indicate that the proportion of teachers who never use more innovative homework strategies, such as assigning problems with no obvious solution method, exceeds the survey reports by a factor of between two to four, depending on the strategy. As a result, the rate of agreement between the two data sources is lowest for the more innovative homework strategies.[22]

CONCLUSIONS

To the extent that we were able to validate the survey data collected on teachers' instructional strategies, we found that such data present an accurate picture of which instructional strategies are used most often by teachers, and they provide some indication of how teachers combine strategies during instruction. Although the picture of teaching

[22]Assignment characteristics for which the direct rate of agreement between the two data sources was 30 percent or lower included: *reading the text or supplementary materials* (30 percent), *applying concepts or principles to different or unfamiliar situations* (13 percent), *solving problems for which there is no obvious method of solution* (30 percent), and *solving applied problems* (28 percent).

that can be drawn from survey data is not a finely-grained one, it is likely to be valid because both the survey and the artifact data clearly show that there is little variation in teachers' instructional strategies. Basically, the majority of teachers use a few instructional strategies and use them often.

Survey data are, however, limited in the precision with which they can measure how frequently teachers use particular strategies. Although teachers may find it easier to respond to questions that provide the five response categories typically included on national surveys, valid distinctions can probably only be made among activities that are done weekly or more often, monthly, once or twice a semester (i.e., infrequently), and never. Our analysis suggests that the distinction between daily activities and those done once or twice a week is not reliable.

Our comparison of teachers' reports on their exam formats, as compared with an analysis of their actual exams, provides an example of where a rather simple re-wording of survey response options can produce more meaningful data. But the other inconsistencies we identified in comparing survey responses with the exams and assignments are symptomatic of a more serious problem. Teachers see their exams and assignments as exhibiting greater variety in their underlying instructional strategies than was evidenced in the artifact coding. Part of the problem might be addressed by providing more precise definitions of what is meant, for example, by problems with *more than one possible approach* or *more than one step to reach a solution*. However, discrepancies between the two types of data sources suggest more serious problems. Teachers see their instruction as more varied and less traditional than is reflected in their exams and assignments, and they do not share common meanings for some of the terms used by curriculum reformers.

The implications for the design of more reliable and valid survey instruments are unclear at this point because so few teachers have adopted the instructional strategies advocated by NCTM and similar groups. Consequently, it is difficult even to conduct valid pilots of alternative survey question wordings or to test new measures of instructional strategies. In the next chapter we attempt to identify more precisely the source of problems and where possible solutions might lie for measuring curriculum

ERIC

during a time of transition. We do so by probing teachers' reports on their instructional goals and then comparing those with the goals reflected in their artifacts and with our analysis of their instructional strategies.

ERIC

Chapter 5

INSTRUCTIONAL GOALS

A final dimension of curriculum consists of the goals or objectives that teachers pursue as they present course content using different instructional strategies. Arguments for including measures of teachers' instructional goals as indicators of curriculum rest on the assumption that the relative emphasis teachers accord different goals reveals something about their choices of instructional strategies. Furthermore, some empirical evidence suggests that teachers using the same textbook emphasize different aspects of it because they value the purposes of instruction differently (for a discussion of why goals should be included as curriculum indicators, see Oakes and Carey, 1989).

However, teachers' reports of their course objectives reflect intended behavior and are less likely to be reliable than reports of actual behavior, such as topic coverage and instructional activities. Despite the obvious problems associated with measuring instructional goals, some have argued that questions about teachers' goals should be included in national surveys because they can function as lead indicators showing the direction in which coursework and teaching in a particular subject may be heading. For example, teachers may report giving some emphasis to goals associated with the mathematics reform movement as a precursor to their engaging in activities consistent with those goals. While in some instances teachers' goals may signal a future change in their behavior, evidence from the implementation of educational innovations suggests that it would be inappropriate to make such an inference in reporting national trends. As McLaughlin (1990) notes in her overview of findings from implementation research, teachers' beliefs may sometimes follow rather than lead their changes in practice, especially if the changes in practice are mandated. So, for example, teachers may be required to integrate topics across different subject areas or have students write in journals, but their belief in the value of those practices may come only after they see that the changes have positive effects on their students.

Our research confirms that instructional goals are the most problematic dimension of curriculum to measure. The consistency between survey responses and instructional artifacts was the lowest among the three dimensions we studied. However, in examining the reasons for the inconsistency, we did learn something about teachers' perceptions and how they integrate new expectations and strategies into their existing approach to teaching. Consequently, we first describe how the teachers in our sample viewed their goals and how the emphasis they reported giving them related to their reported use of instructional strategies. We then examine the consistency of teachers' self-reports with what coders identified as the goals reflected in teachers' exams and assignments.

IDENTIFYING TEACHERS' INSTRUCTIONAL GOALS FROM SURVEY DATA: SOME ILLUSTRATIVE EXAMPLES

Teachers were asked to rate the emphasis they gave to twenty different instructional goals. Figures 5.1 and 5.2 indicate that although teachers' emphasis on goals that might be considered more traditional was somewhat greater, a majority also reported giving either a moderate or major emphasis to most of the reform goals listed.[23]

As we did with instructional activities, we grouped together those instructional goals associated with the mathematics reform movement and another set that could be considered more traditional. Those scales are displayed in Table 5.1. The reform goals scaled well, but the traditional goals did not. We can only speculate that the conventional goals did not scale well because, unlike the reform goals, they are not based on a coherent theory of instruction. Rather, they represent a set of goals that teachers have traditionally pursued, perhaps without regard to the linkages among them.

We also conducted a factor analysis to determine if there were some subsets of goals that were related to each other in a meaningful way. Four factors emerged that could be substantively interpreted. The first factor shown in Table 5.2 includes six items that all deal with students developing critical thinking skills. The second factor includes four items that stress having students understand mathematical relationships in different

[23]A majority of respondents reported giving a moderate or major emphasis to *learning to represent problem structures in multiple ways* (79 percent), *integrating different branches of mathematics* (76 percent), *raising questions and formulating conjectures* (77 percent), *finding examples and counterexamples* (66 percent), *judging the validity of arguments* (51 percent), and *discovering generalizations* (69 percent), in addition to the three reform goals displayed in Figure 5.1.

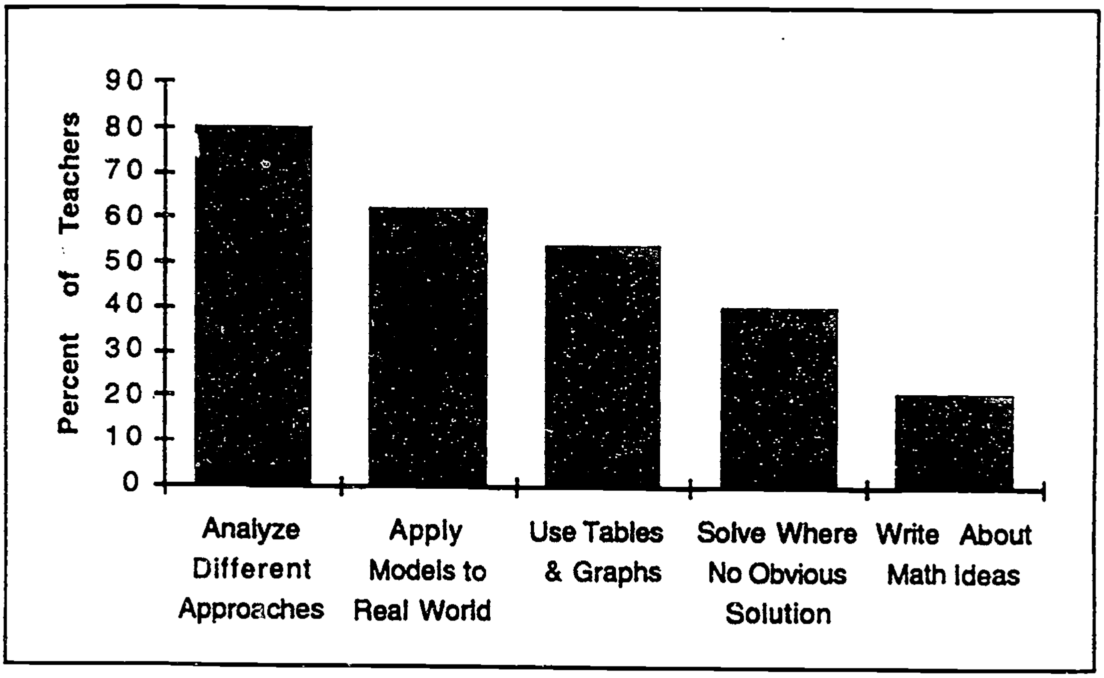

Figure 5.1 -- Proportion of Teachers Reporting Major or Moderate Emphasis on Selected Reform Goals

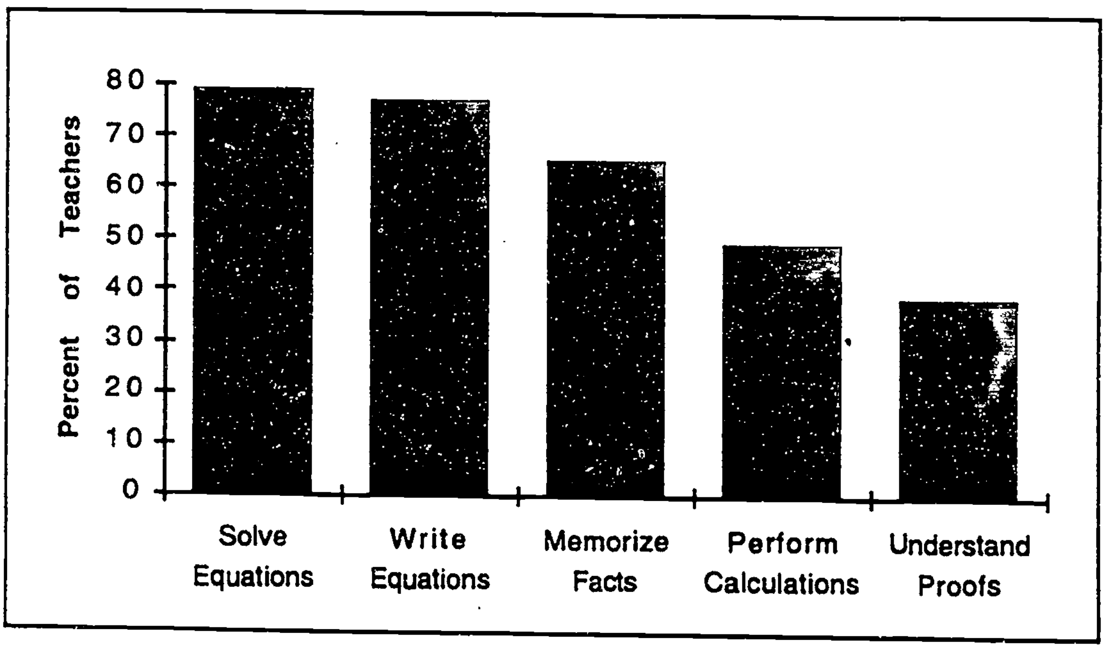

Figure 5.2 -- Proportion of Teachers Reporting Major or Moderate Emphasis on Traditional Goals

BEST COPY AVAILABLE

Table 5.1
Instructional Goal Scales

	Reform	Traditional
Understanding the nature of proof		+
Memorizing facts, rules and steps		+
Learning to represent problem structures in multiple ways (e.g., graphically, algebraically, numerically)	+	
Integrating different branchs of mathematics (e.g., algebra, geometry) into a unified framework	+	
Conceiving and analyzing the effectiveness of different approaches to problem solving	+	
Performing calculations with speed and accuracy		+
Showing the importance of math in daily life		
Solving equations		+
Raising questions and formulating conjectures	+	
Increasing students' interest in math		
Integrating math with other subjects	+	
Finding examples and counterexamples	+	
Judging the validity of arguments	+	
Discovering generalizations	+	
Representing and analyzing relationships using tables, charts and graphs	+	
Applying mathematical models to real-world phenomena	+	
Writing about mathematical ideas	+	
Designing a study or experiment	+	
Writing equations to represent relationships		+
Solving problems for which there is no obvious method of solution	+	
	α = .86	.37

Table 5.2
Instructional Goals Factor Matrix*

	Factor 1	Factor 2	Factor 3	Factor 4
Raising questions and formulating conjectures	.81	.12	-.14	.09
Judging the validity of arguments	.73	.07	-.11	.21
Understand the nature of proof	.72	-.05	-.06	-.10
Increasing students' interest in math	.60	.24	-.11	.13
Finding examples and counterexamples	.59	.13	.22	.14
Discovering generalizations	.59	.34	-.06	.23
Learning to represent problem structures in multiple ways (e.g., graphically, algebraically, numerically)	.17	.72	.03	.12
Conceiving and analyzing the effectiveness of different approaches to problem-solving	.41	.57	-.02	-.15
Writing equations to represent relationships	-.08	.56	.06	.21
Integrating different branches of mathematics (e.g., algebra, geometry) into a unified framework	.35	.47	-.28	-.05
Solving equations	-.05	.30	.64	.15
Writing about mathematical ideas	.12	.23	-.61	.29
Solving problems for which there is no obvious method of solution	.16	.27	-.53	-.04
Applying mathematical models to real world phenomena	.21	.15	-.05	.70
Showing the importance of math in daily life	-.04	.08	.17	.52
Integrating math with other subjects	.35	-.02	-.04	.47
Designing a study or an experiment	.23	-.07	-.44	.47
Memorizing facts, rules, and steps	-.26	-.26	.04	-.01
Performing calculations with speed and accuracy	.18	.15	.11	.03
Representing and analyzing relationships using tables, charts, and graphs	.07	.42	-.30	.31
Eigenvalue	4.82	1.48	1.43	1.20

* Although five factors were extracted, the four primary factors are presented here because the scree diagram indicated that the remaining factor did not explain a substantial, additional proportion of the variance, and it was not readily interpretable on substantive grounds.

ERIC

ways. Like the first two factors, the fourth is consistent with the goals of the mathematics reform movement and deals with the application of mathematics to other subjects and to daily life. Only the third factor contains items that might be considered more traditional: teachers who emphasize solving equations give little emphasis to writing about mathematical ideas and to solving problems for which there is no obvious method of solution.

We expected to see a positive correlation between reform goals and reform instructional activities and between traditional goals and traditional modes of teaching. The correlation between the reform goal scale and the reform instructional repertoire in Table 4.1 is strongly positive (r=.75), while the correlation between reform goals and the traditional teaching repertoire is negative (r=-.51). However, traditional goals and instructional activities are not correlated. The correlation between the traditional goal scale and the scale of traditional instructional activities is negative (r=-.18) and nonsignificant. The major reason is the lack of variation on these two scales: two-thirds of the teachers reported a major or moderate emphasis on three or more of the five traditional goals and 71 percent reported engaging in six or more of the nine traditional instructional strategies at least once or twice a week.

Our data confirm what Cohen and Peterson (1990) found in their study of the California mathematics framework--viz., that even teachers who endorse curriculum reform and implement it in their own classrooms do so by integrating the new with the traditional. Although close to half of the teachers in our sample (46 percent) report that they emphasize most of the reform goals in their teaching, only 12 percent of the sample engage in four or more reform instructional activities at least once or twice a week. However, we did identify seven teachers (10 percent of the sample) who reported a moderate or major emphasis on nine or more reform goals and who also reported using four or more reform-oriented instructional strategies at least once or twice or week. However, all but two of these seven teachers also use at least half of the traditional instructional strategies just as frequently. In other words, by their own self-reports, few respondents in our sample rely on the instructional strategies that mathematics reformers espouse for advancing reform goals these teachers seem to accept. Furthermore, even the few respondents who might--by their own reports--be characterized as "reform teachers" still use traditional teaching strategies as part of their instructional repertoires. Consistent with the implementation patterns that characterize the adoption of many classroom innovations, these teachers are layering new practices onto their existing ones.

CONSISTENCY BETWEEN THE SURVEYS AND THE ARTIFACTS

The difficulty in interpreting data on instructional goals is further confounded when we compare teachers' self-reports on the survey with the artifact data coded from their assignments and exams. Figure 5.3 shows the rate of agreement between the surveys and exams and between the surveys and assignments on the degree of emphasis that teachers gave each of the instructional goals. The level of consistency between teachers' self-reports on the surveys and coders' depiction of their teaching gleaned from the artifacts was considerably less for goals than for either topic coverage or instructional strategies. However, the rate of agreement was slightly higher for assignments than for exams, perhaps because there were more data points from which to make inferences. On the whole, survey and artifact data were more consistent for traditional goals (four of the five traditional goals had rates of agreement above the mean for the entire list) than for reform goals (four of the 13 reform goals were above the mean).

The major source of the discrepancies could be traced to the coders' very different judgments about the amount of emphasis that teachers were giving reform goals. For 12 of the 13 reform goals, coders indicated that 75 percent of the teachers had given these instructional objectives either a minor or no emphasis. This depiction of teachers' goals is generally consistent with the picture of their exams and assignments that they themselves provided in their survey responses. For example, the survey data presented in Figure 4.2 indicate that only a small proportion of teachers' exams require students to describe how to solve problems, explain their reasoning, or apply concepts to unfamiliar situations. On the other hand, when asked to characterize their instruction through the lens of the goals they stress, teachers presented a very different picture. For only two of the reform goals (*writing about mathematical ideas, designing a study or experiment*) did an equally high proportion of teachers report a small emphasis, thus agreeing with the coders. As noted previously, close to a majority reported giving a

ERIC

Table 5.3
Instructional Goals: Consistency between Surveys and Exams and between Surveys and Assignments

	Exams		Assignments	
	% Direct Agreement	% Within One Survey Response Category	% Direct Agreement	% Within One Survey Response Category
Understanding the nature of proof	36.6	70	37.5	79.7
Memorizing facts, rules, and steps	26.2	78.7	29.7	79.7
Learning to represent problem structures in multiple ways (e.g., graphically, algebraically, numerically)	13.1	49.2	20.3	60.9
Integrating different branches of mathematics (e.g., algebra, geometry) into a unified framework	14.8	39.3	21.9	53.1
Conceiving and analyzing the effectiveness of different approaches to problem solving	3.3	26.2	7.8	32.8
Performing calculations with speed and accuracy	30.2	87.3	30.2	87.3
Showing the importance of math in daily life	6.6	31.2	17.2	48.4
Solving equations	24.6	64	29.7	78.1
Raising questions and formulating conjectures	5.8	28.9	11.5	42.3
Increasing students' interest in math	5	18	9.4	37.5
Integrating math with other subjects	6.7	48.3	18	63.9
Finding examples and counterexamples	9.8	36.1	11.1	49.2
Judging the validity of arguments	21.7	50	25.8	58.1
Discovering generalizations	9.8	34.4	14.3	42.9
Representing and analyzing relationships using tables, charts, and graphs	23.3	61.7	29	71
Applying mathematical models to real-world phenomena	9.6	50	11.5	51.9
Writing about mathematical ideas	41.7	81.7	43.5	88.7
Designing a study or experiment	53.3	90	58	91.7
Writing equations to represent relationships	14.8	54	19.4	58
Solving problems for which there is no obvious method of solution	17.3	63.5	21.2	65.4
Mean	19	53	23.4	62

moderate or major emphasis to most reform goals and for four goals, 75 percent or more reported doing so.[24]

It was these patterns--low levels of agreement between the survey and artifact data, more problems with reform than traditional goals, and teachers reporting a greater emphasis on reform goals than coders could detect--that initially led us to re-interview a subsample of teachers. One problem we discovered in these follow-up interviews is the different ways that teachers interpreted the response options (*major, moderate, minor,* and *none*) for the goals item. This same set was used on the NELS-SFU questionnaire and a variant of it has also been used on NAEP teacher surveys. But teachers interpret this response option quite differently. Some assumed that the underlying dimension was the frequency with which they undertook activities consistent with a particular goal, while others assumed that emphasis should be defined in terms of how important they considered a goal for their students' understanding, regardless of how often they undertook activities reflective of that goal. Other teachers combined frequency and importance in their assessment of emphasis.

Coders were instructed to base their judgments on the prevalence of tasks consistent with a particular goal--for reform goals, those tasks were identified from NCTM materials--and the goal's relative importance as compared with other objectives the teacher seemed to be stressing. The notion that some teachers might place a major emphasis on a goal but not incorporate it into many activities--e.g., by stressing it with great clarity and forcefulness at a few key points during the course--is not something that we could measure well with artifacts.

A second, and by far greater, problem is the different meanings that teachers ascribed to terms associated with the mathematics reform movement. Table 5.4 illustrates differing interpretations of a reform goal which had the lowest level of agreement between the survey and the exam and assignment artifacts. At a general level, five of the six teachers interpreted the goal in a way consistent with its reform

[24]The four goals for which 75 percent or more of the teachers reported giving them a moderate or major emphasis are displayed either in Figure 5.1 or listed in footnote 23.

Table 5.4
Examples of Teachers' Interpretations of A Reform Goal

Item	Reform Definition	Teacher Interpretations*
Conceiving and analyzing the effectiveness of different approaches to problem-solving	Helping students to be able to think of problem-solving strategies such as trial and error and other iterative methods, and to judge when to use a particular strategy Apply such strategies in solving "real world" and non-routine problems.	• When we're working with signed numbers, I tell students you should go by the rules, but if you forget the rules, you can always figure out the answer logically. There are two methods: going by the rules or doing it logically." (General math teacher) • "I do a lot of statement problems. I draw on the board." (Pre-algebra teacher) • "In Math A, we don't have too many different approaches. When I first read the question, I saw it asking about student approaches. Now I see it as asking about what the teacher does." (Pre-algebra teacher) • "This goal is about how much discussion and presentation you do about alternative methods of solution. I don't consciously think of it; it naturally comes up." (Geometry teacher) • "This is more of a classroom activity where you talk about different approaches, especially if a student asks, 'Could I have solved the problem this way?'" (Math analysis teacher) • When we come to new problems, I don't tell students anything, hoping they will ask good questions. Then I give them the background--the classical approach. I always encourage creative thinking, different ways to solve problems." (AP calculus teacher)

• In the group interviews, teachers were asked to tell us what types of activities they saw as representing a particular goal in the course they reported on in their survey.

meaning--i.e., encouraging more than one solution method. But only the calculus teacher's discussion comes close to the notion of "conceiving," and most of the teachers seem to be interpreting problem-solving in a narrower sense of solving traditional mathematics problems, rather than strategies for solving "real world" or non-routine problems.

This example of disparate interpretations is by no means unique. The previous chapters reported problems with other reform-oriented terms. Not only did teachers have differing interpretations of these terms, but in a number of cases they reported not knowing at all what the phrases meant.

CONCLUSIONS

Our analysis suggests that instructional goals are too problematic to be validly measured through national surveys of teachers. The data are inconsistent not only with artifact data, but also with teachers' own self-reports on other survey items such as those describing their exam formats. These inconsistencies between teachers' reports about their goal emphasis and their instructional strategies are difficult to interpret. It might be that the lack of a consistent relationship stems from the different meanings teachers ascribe to terms associated with the mathematics reform movement. Or, it may be that acknowledging the importance of particular goals is a precursor to implementing instructional practices consistent with those objectives. Or, it might be that despite teachers' willingness to report candidly about their reliance on traditional instructional strategies, social desirability becomes a factor in talking about their philosophy of teaching. These are among a number of plausible explanations for the disjuncture between teachers' reported goals and classroom practice. However, at this point we do not know which actually account for the inconsistencies. As a result, survey data on instructional goals cannot be unambiguously interpreted.

Consequently, we would recommend that questions about teachers' instructional goals be deleted from national surveys. These items could then be replaced with more detailed measures of topic coverage--thus improving the amount and quality of data on the most central aspect of curriculum, without greatly increasing respondent burden. At least in the short-term, data on teachers' goals might be more effectively gathered through smaller, supplemental studies. They might be collected as part of a validation study so that teachers' self-reports could be compared with their instructional artifacts; data might be collected using face-to-face, open-ended interviews, perhaps combined with classroom observations; or focus group and similar strategies might be used to probe the meanings that teachers ascribe to different goals. Interpreting survey data about attitudes and beliefs is always difficult, but in the case of teachers' goals, the dangers of misinterpretation seem particularly high and appear to outweigh the value of obtaining information through a relatively inexpensive, broad-based method.

ERIC

Chapter 6

DESIGN CHOICES FOR IMPROVED CURRICULUM INDICATORS

As curriculum assumes greater prominence on the education policy agenda, the demand for better indicators will continue. As a result, three questions face those responsible for the design and operation of educational indicator systems:

- how will curriculum indicator data be used,
- how much do various users need to know about curriculum, and
- what is the most effective design for collecting curriculum indicator data?

These are difficult questions to answer, and judgments about appropriate directions will be shaped as much by political values and resource constraints as by technical considerations. Nevertheless, the findings and conclusions from this project can help inform the decision process.

USES

The potential uses of curriculum indicator data could conceivably range from the kind of national snapshot now provided by NAEP, NELS, and other similar national surveys to the high stakes applications implied in some proposed uses of opportunity-to-learn standards (McDonnell, 1995). We would argue that an enhanced version of existing surveys will provide a reasonably valid depiction of the mathematics curriculum in this country. Nevertheless, there will be two major limitations: the characterization will be a rather general one and it may not provide a very accurate picture of either teachers' intentions or practices with regard to curriculum reform. Still, it is possible to obtain sound information about the depth and breadth of course content, how it varies across courses and types of schools, and a better indication of teachers' instructional repertoires than is currently available.

However, despite the improvements that can be made in surveys over the next several years, we do not believe that the information collected will meet the necessary criteria for high stakes uses. The data will still be at such a level of generality that they cannot be used to make valid determinations about the alignment of individual schools with any type of content standards. Yet due process would require that valid and reliable measures of each school's curriculum be established before it could be held accountable for its instructional activities. Given the measurement and interpretation problems we have identified, we do not believe that curriculum indicator data could meet such a legal standard in the near future.[25] Therefore, the most appropriate uses will continue to be informational ones. Curriculum indicator data can provide a general picture of the distribution of OTL across different types of schools and students and it can chart overall trends in curricular practice, but it cannot serve as the basis for decisions with potentially serious consequences for schools and teachers.

INFORMATION NEEDS

The second question raises the issue of what should be included in the domain of curriculum. In our study, we focused on content coverage, instructional strategies, and goals, and most indicator designers recommend some variant of these three categories. However, these categories are largely teacher-centered, and do not directly measure the role of students in constructing knowledge. Measuring active student learning greatly complicates both the measurement and the data collection task, and would likely necessitate more data than can be obtained from teacher surveys. However, it may soon be possible to consider another potentially large data base as a source for curriculum indicators. The increased use of student portfolios by states and local districts provides an opportunity to experiment with using them not just as the basis for assessing students, but also as sources of information about the nature of the teaching and learning process. Up to this point, research on student portfolios has focused on scoring them as measures of student achievement, but a parallel development effort could focus on how to extract data that might serve as indicators of the types of instructional strategies being used and students' role in those activities.

Even if the curriculum is defined more narrowly in terms of the three categories

[25]Another issue that would arise if curriculum data were collected for high stakes purposes relates to the quality of teachers' survey responses. We found few social desirability problems in their responses. However, our surveys were administered under very low stakes conditions. All the research showing that teachers change their behavior in response to the content and format of student assessments strongly suggests that under high stakes conditions, teachers would likely bias their responses. They might find it in their interest to report responses consistent with policymakers' expectations, thus corrupting the information collected. As a result, validation studies would need to be conducted much more frequently than if the data were only for informational purposes and no direct consequences for teachers were attached to its use.

we used and confined to indicators that can be effectively measured through teacher surveys, the level of detail desired within each of these categories can vary considerably. Given the relationship between students' curricular exposure and their achievement, as well as our study results showing that surveys can provide reasonably accurate measures of topic coverage, we recommend that future national surveys place a greater emphasis on topic coverage. Not only are the topics currently included on national teacher surveys too few and too general to provide a valid picture of OTL, the information they generate is virtually useless in understanding curricular trends. Future items on topic coverage should be tailored to specific course levels, and should include more topics at a greater level of specificity. Our post-data collection questionnaire is an example of such an enhanced survey.

Although we would accord it lower priority, we also recommend including a more comprehensive set of items dealing with instructional strategies. Asking teachers about a broader range of classroom practices would provide better information about the different ways that they combine strategies and how they integrate newer practices into their traditional repertoires. The findings from our study and a number of others indicate that teachers rely on only a few traditional strategies. Yet the expectation continues that they will adopt a variety of instructional reforms. Whether that expectation is met or not remains an open question. But asking teachers about only a few traditional and a few reform practices ignores the reality of policy implementation. If teachers do adopt the instructional strategies advocated by reformers, it will be through a process of adaptation and layering (Darling-Hammond, 1990). Without a fairly comprehensive set of instructional practice items, it will be difficult to determine exactly what these hybrid repertoires look like or how consistent they are with reformist guidelines.

We recognize that our recommendations would require additional time for survey administration and hence increase respondent burden. This trade-off between improved data quality and respondent burden is a particular problem in the case of national surveys used to collect a variety of different data from the same respondents. However, as we argued in the previous chapter, teachers' instructional goals cannot be validly measured through survey data. Therefore, the additional burden associated with an enhanced survey on topic coverage and instructional strategies could be reduced somewhat by eliminating those items dealing with instructional goals.

DATA COLLECTION STRATEGIES

Decisions about use and scope will largely determine data collection strategies. Our findings suggest three areas of possible investment. The first has already been discussed: improving the design of national surveys. In addition to changing the relative emphasis accorded different aspects of curriculum, a number of suggested changes in item wordings and response option scales were outlined in previous chapters. These changes can be implemented quite cost-efficiently.

A second area of future investment are indepth studies on small samples of teachers and classrooms to monitor changes in mathematics teaching. These studies would use techniques that can measure instructional processes with greater subtlety than is possible through surveys. The more complete, nuanced data about such issues as teachers' understanding of reform goals and their different uses of reform strategies can then be used to interpret survey results and to improve the design of future surveys.

The final area for future investment is the one that has been the primary focus of this study. We believe that the kind of validation study we have piloted should be integrated into the design of curriculum indicator systems. The primary, and most pressing, reason for such validation studies is the current reform context. Proposed changes in curriculum content and instructional practice mean that the language of mathematics teaching is in flux, and teachers do not share a common understanding of key terms. The effect is likely to be either a serious misinterpretation of survey results or an inability to interpret them at all. The solution is to make problematic survey items clearer through the use of more precise definitions and concrete examples. However, as we noted in the case of instructional strategies, so few teachers have adopted the new approaches that it is difficult to test alternative survey question wordings or experiment with new measures. Consequently, until language and practice have stabilized, validation studies (perhaps combined with indepth case studies and focus group interviews) will need to be an integral part of curriculum indicator systems.

Although current interest in curriculum reform and hope for its widespread implementation provide the primary rationale for validation studies, they would still be needed even in more stable times. By collecting detailed data from multiple sources over shorter periods of time, validation studies can provide benchmarks against which to judge both the validity and reliability of survey data. It is only with such data that we can know whether teachers are reporting reliable estimates of topic coverage or whether their characterizations of exams and assignments are accurate. Such independently-collected information helps not only in interpreting survey data, but also identifies sources of measurement error and informs the design of future surveys.

However, validation studies do not have to be conducted every time a national survey is administered. Rather, we would recommend conducting one only when a new survey effort is begun--e.g., at the beginning of a longitudinal study like NELS or when major design changes are implemented in the NAEP teacher survey. The validation study would then be conducted as part of the first administration of the survey, with such efforts required only every five years or more.

Although we would recommend several modifications in the procedures used in our pilot study, we believe that the basic structure is sound. The instructional artifacts worked well as benchmarks and despite some obvious limitations, were easily collected from teachers. Although coding artifacts to extract information comparable to that collected from the surveys was a difficult task, we now have a template that can be improved upon and replicated quite easily. Given what we have learned from the pilot study, we are confident that the level of inter-rater agreement can be increased. The coding specifications can now be made more precise, and the coding process organized so that coders' work is reviewed more frequently through a moderation process that identifies discrepant judgments and makes appropriate adjustments. The coding of instructional artifacts will never be as reliable as, for example, the scoring of open-ended test items because the type and mix of material is unstandardized across teachers. Nevertheless, we believe that by using the survey categories as the basis for a content analysis of the artifacts and by closely monitoring the coding process, high quality benchmark data can be obtained.

ERIC

In order to make valid comparisons across courses, future validation samples will need to be somewhat larger--probably about twice as large as for the pilot study. However, given that there is less variation in the curriculum of upper-level course such as calculus and that policy concerns about opportunity-to-learn are greatest in lower-level courses, one option might be to concentrate the study's focus on courses at or below algebra II. A particular emphasis might be on lower level courses such as pre-algebra and on those that integrate topics across traditional course categories.

The similarity in our findings about teachers' instructional practices with those from larger, nationally-representative samples suggests that our smaller sample is generally reflective of high school mathematics teaching. However, in order to avoid idiosyncracies that might characterize the teacher force in one or two states, future validation studies should include teachers from a larger number of states. For example, the proportion of California mathematics teachers who have a college major in mathematics is considerably below the national average (44 percent in 1991, as compared with a national average of 69 percent) (Blank and Gruebel, 1993).[26] With the modifications outlined, the basic approach used in this pilot study should serve as an effective template for future validation studies.

Over the past decade, the quality of education indicators has steadily improved, with the greatest progress made in indicators of school and classroom processes. The "black box" that characterized older input-output models has been replaced with an increasingly comprehensive set of indicators that can report national trends in school organization and curriculum. But the failure to validate these indicators has remained a problem. Because items are typically transferred from one survey to another with no attempts at validation, the extent to which they measure how students are actually taught was virtually unknown. This study represents a first step in ensuring that curriculum indicators are valid and reliable measures of the instruction occurring in the nation's classrooms.

---

[26]We focused on California because we assumed that the state's innovative curriculum frameworks would mean that more reform-oriented teachers would be included in our sample. However, like many others, we underestimated how difficult and slow implementation of the frameworks would be.

REFERENCES

Blank, R. K., & Gruebel, D. (1993). *State indicators of science and mathematics education.* Washington, DC: Council of Chief State School Officers.

California Department of Education. (1992). *Mathematics framework for California public schools.* Sacramento, CA: Author.

Cohen, D. K., & Peterson, P. L. (1990). Special issue of *Educational Evaluation and Policy Analysis, 12* (3), 233-353.

Darling-Hammond, L. (1990). Instructional policy into practice: "The power of the bottom over the top." *Educational Evaluation and Policy Analysis, 12* (3), 339-347.

Freeman, D. J., Kuhs, T. M., Porter, A. C., Floden, R. E., Schmidt, W. H. & Schwille, J. R. (1983). Do textbooks and tests define an national curriculum in elementary school mathematics? *Elementary School Journal, 83* (5), 501-513.

Goodling, B. (1994, March 23). A failed policy about inputs. *Education Week,* XIII(26), 36.

Harp, L. (1994, May 18). The plot thickens: The real drama behind the Kentucky education reform act may have just begun. *Education Week,* pp. 20-25.

Jones, L. V., Davenport, E. C., Bryson, A., Bekhuis, T., & Zwick, R. (1986). Mathematics and science test scores as related to courses taken in high school and other factors. *Journal of Educational Measurement, 23* (3), 197-208.

Kifer, E. (1993). Opportunities, talents and participation. In L. Burstein (Ed.), *The IEA study of mathematics III: Student growth and classroom processes* (pp. 279-307). New York: Pergamon.

McDonnell, L. M., Burstein, L., Ormseth, T., Catterall, J. M., & Moody, D. (1990). *Discovering what schools really teach: Designing improved coursework indicators.* Santa Monica, CA: RAND.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis.*

McLaughlin, M. W. (1990). The RAND Change Agent Study revisited: Macro perspectives and micro realities. *Educational Researcher,* 19(9), 11-16.

McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., & Cooney, T. J. (1987). *The underachieving curriculum: Assessing U.S. school mathematics from an international perspective*. Champaign, IL: Stipes Publishing.

Merl, J. (1994, May 6). Furor continues to build over state's CLAS exams. *Los Angeles Times*, pp. A1, 18.

Mullis, I. V. S, Dossey, J. A., Owen, E. H., & Phillips, G. W. (1991). *The state of mathematics achievement: NAEP's 1990 assessment of the nation and the trial assessment of the states* (OERI Rep. No. 21-ST-04). Princeton, NJ: Educational Testing Service.

Mullis, I. V. S, Jenkins, F., & Johnson, E. (1994). *Effective schools in mathematics: Perspectives from the NAEP 1992 assessment* (OERI Rep. No. 23-RR-01). Washington, D.C.: U.S. Government Printing Office.

Murnane, R. J., & Raizen, S. A. (1988). *Improving indicators of the quality of science and mathematics education in grades K-12*. Washington, DC: National Academy Press.

National Council on Education Standards and Testing. (1992). *Raising standards for American education*. Washington, DC: U.S. Government Printing Office.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.

Oakes, J. & Carey, N. (1989). Curriculum. In R.J. Shavelson, L.M. McDonnell, J. Oakes (Ed.) *Indicators for monitoring mathematics and science education: A sourcebook*. Santa Monica, CA: RAND.

National Study Panel on Education Indicators. (1991). *Education counts: An indicator system to monitor the nation's educational health*. Washington, DC: U.S. Government Printing Office.

O'Day, J. A., & Smith, M. S. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent education policy* (pp. 250-313). San Francisco: Jossey-Bass.

OERI State Accountability Study Group. (1988). *Creating responsible and responsive accountability systems*. Washington, DC: U.S. Department of Education.

Owens, M. R. (1994, March 23). The name of the camel is truth. *Education Week, XIII(26)*, 35-36.

Porter, A. C. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis, 13(1)*, 13-29.

Porter, A. C., Kirst, M. W., Osthoff, E. J., Smithson, J. L., & Schneider, S. A. (1993). *Reform up close: An analysis of high school mathematics and science classrooms*. University of Wisconsin-Madison, Wisconsin Center for Education Research, School of Education.

Raizen, S. A., & Jones, L. V. (Eds.). (1985). *Indicators of precollege education in science and mathematics: A preliminary review*. Washington, DC: National Academy Press.

Ravitch, D. (1995). *National standards in American education*. Washington, DC: Brookings.

Rothman, R. (1993, April 7). 'Delivery' standards for schools at heart of new policy debate. *Education Week*, p. 21.

Schmidt, W. H., Wolfe, R. G., & Kifer, E. (1993). The identification and description of student growth in mathematics achievement. In L. Burstein (Ed.), *The IEA study of mathematics III: Student growth and classroom processes* (pp. 59-75). New York: Pergamon.

Shavelson, R., McDonnell, L. M., Oakes, J., & Carey, N. with Picus, L. (1987) *Indicator systems for monitoring mathematics and science education*. Santa Monica, CA: RAND.

Shavelson, R. J., McDonnell, L. M., & Oakes, J. (Eds.). (1989) *Indicators for monitoring mathematics and science education: A sourcebook*. Santa Monica, CA: RAND.

Travers, K. J., Garden, R. A., & Rosier, M. (1988). Introduction to the study. In D. F. Robitaille and R. A. Garden (Eds.), *The IEA study of mathematics II: Contexts and outcomes of school mathematics* (pp. 1-16) New York: Pergamon.

Travers, K. J., (1993). Overview of the longitudinal version of the second international mathematics study. In L. Burstein (Vol. Ed.), *The IEA study of mathematics III: Student growth and classroom processes*. (pp. 1-27). New York: Pergamon.

Travers, K. J. and Westbury, I. (Eds), *The IEA study of mathematics I: Analysis of mathematics curricula* (1989). New York: Pergamon.

Weiss, I. (1994). *A profile of science and mathematics education in the United States: 1993*. Chapel Hill, NC: Horizon Research, Inc.

Wittrock, M. C. (Ed.). (1985). *Handbook of research on teaching* (3rd ed.). New York: Macmillan.

APPENDIX

Initial Teacher Survey (administered prior to artifact data collection)

Enhanced Teacher Survey (administered after artifact data collection)

Daily Log Form

VALIDATING NATIONAL CURRICULUM INDICATORS

INITIAL TEACHER SURVEY

This questionnaire asks for some initial information about the goals, content, and instructional activities in the class that has been chosen for the RAND/UCLA study on validating curriculum indicators. This information, along with the instructional materials you will be providing, will help in describing students' educational experiences.

The survey includes questions about characteristics of the class, teaching strategies, curriculum content, and general information about your teaching experience.

Please mark your responses directly on the questionnaire. Place it in the envelope with your class assignments for the first week, and return it to RAND.

THANK YOU FOR YOUR CONTRIBUTION TO THIS STUDY.

Class Information

1. Identity Code: ______

2. Class Title: ______________________________

3. How many students are enrolled in this class?

No. of Students: ☐☐

4. How many students in this class are from minority racial/ethnic groups (e.g., Black, Hispanic, Asian)? (If unsure, give your best estimate.)

No. of Students: ☐☐

5. Which of the following best describes the level this class is considered to be?

(Circle One)

Remedial	1
General	2
Voc/Tech/Business	3
College Prep/Honors	4
AP	5

6. Which of the following best describes the achievement level of the students in this class compared with the average student in this school?

(Circle One)

Higher achievement levels	1
Average achievement levels	2
Lower achievement levels	3
Widely differing achievement levels	4

7. Approximately how much homework do you typically assign each day to this class?

Minutes: ☐☐

8. How often do you do each of the following with homework assignments?

(Circle One Number on Each Line)

	Never	Some of the Time	Most of the Time	All of the Time
a. Keep records of who turned in the assignment	1	2	3	4
b. Return assignments with grades or corrections	1	2	3	4
c. Discuss the completed assignment in class	1	2	3	4

9. Approximately how many minutes per week does this class meet regularly (not including lab periods)?

Minutes: ☐☐

10. Approximately how may minutes per week does this class have lab sessions? (If there is no lab, enter "00.")

Minutes: ☐☐

11. Indicate about what percent of class time is spent in a typical week doing each of the following with this class.

(Circle One Number on Each Line)

	None	≤10%	10-24%	25-49%	50-74%	75-100%
a. Providing instruction to the class as a whole	1	2	3	4	5	6
b. Providing instruction to small groups of students	1	2	3	4	5	6
c. Providing instruction to individual students	1	2	3	4	5	6
d. Maintaining order/disciplining students	1	2	3	4	5	6
e. Administering tests or quizzes	1	2	3	4	5	6
f. Performing routine administrative tasks (e.g., taking attendance, making announcements, etc.)	1	2	3	4	5	6
g. Conducting lab periods	1	2	3	4	5	6

12. **How often do you use the following teaching methods or media?**

(Circle One Number on Each Line)

	Never/ Barely	1-2 Times a Month	1-2 Times a Week	Almost Everyday	Everyday
a. Lecture	1	2	3	4	5
b. Use computers	1	2	3	4	5
c. Use audio-visual material	1	2	3	4	5
d. Have teacher-led whole-group discussion	1	2	3	4	5
e. Have students respond orally to questions	1	2	3	4	5
f. Have student-led whole-group discussions	1	2	3	4	5
g. Have students work together in cooperative groups	1	2	3	4	5
h. Have students complete individual written work	1	2	3	4	5
i. Have students give oral reports	1	2	3	4	5

13. **Indicate the importance you give to each of the following in setting grades for students in your classes (excluding special education students).**

(Circle One Number on Each Line)

	Not Important	Somewhat Important	Very Important
a. Achievement relative to the rest of the class	1	2	3
b. Absolute level of achievement	1	2	3
c. Individual improvement or progress over past performance	1	2	3
d. Effort	1	2	3
e. Class participation	1	2	3
f. Completing homework assignments	1	2	3
g. Consistently attending class	1	2	3

For Math Teachers Only

Those teaching science classes should *SKIP TO QUESTION 16* on the following page.

14. In this math class, how much emphasis do you give to each of the following objectives?

(Circle One Number on Each Line)

	None	Minor	Moderate	Major
a. Understanding the nature of proofs	1	2	3	4
b. Memorizing facts, rules, and steps	1	2	3	4
c. Learning to represent problem structures in multiple ways (e.g. graphically, algebraically, numerically, etc.)	1	2	3	4
d. Integrating different branches of math (e.g., geometry, algebra) into a unified framework	1	2	3	4
e. Conceiving and analyzing effectiveness of multiple approaches to problem solving	1	2	3	4
f. Performing calculations with speed and accuracy	1	2	3	4
g. Showing importance of math in daily life	1	2	3	4
h. Solving equations	1	2	3	4
i. Raising questions and formulating conjectures	1	2	3	4
j. Increasing students' interest in math	1	2	3	4

15. Have you taught or reviewed the following topics in this math class during this year? (If you have reviewed and taught an item as new content, mark #3 only.)

(Circle One Number on Each Line)

	No, but it was was taught previously	Yes, but I reviewed it only	Yes, but I taught it as new content	No, but I will teach or review it later this year	No, topic is beyond the scope of this course
a. Integers	1	2	3	4	5
b. Patterns and functions	1	2	3	4	5
c. Linear Equations	1	2	3	4	5
d. Polynomials	1	2	3	4	5
e. Properties of geometric figures	1	2	3	4	5
f. Coordinate Geometry	1	2	3	4	5
g. Proofs	1	2	3	4	5
h. Trigonometry	1	2	3	4	5
i. Statistics	1	2	3	4	5
j. Probability	1	2	3	4	5
k. Calculus	1	2	3	4	5

ERIC

For Science Teachers Only

Those teaching math classes only should *SKIP TO THE SECTION MARKED* **Teacher Background**

16. In this science class, how much emphasis do you give to the following objectives?

(Circle One Number on Each Line)

	None	Minor	Moderate	Major
a. Increasing students' interest in science	1	2	3	4
b. Learning and memorizing scientific facts, principles, and rules	1	2	3	4
c. Learning scientific methods	1	2	3	4
d. Preparing students for future study in science	1	2	3	4
e. Developing problem solving/inquiry skills	1	2	3	4
f. Developing skills in lab techniques	1	2	3	4
g. Learning about applications of science to environmental issues	1	2	3	4
h. Showing importance of science in daily life	1	2	3	4

17. How often do you do each of the following activities in this science class?

(Circle One Number on Each Line)

	Never/ Barely	1-2 Times a Month	1-2 Times a Week	Almost Everyday	Everyday
a. Have students do an experiment or observation individually or in small groups	1	2	3	4	5
b Demonstrate an experiment or lead students in systematic observations	1	2	3	4	5
c. Require students to turn in written reports on experiments or observations	1	2	3	4	5
d. Discuss current issues and events in science	1	2	3	4	5
e. Have students use computers for data collection and analysis	1	2	3	4	5
f. Use computers for demonstrations/simulations	1	2	3	4	5
g. Have students give oral reports	1	2	3	4	5
h. Have students independently design and conduct their own science projects	1	2	3	4	5
i. Discuss career opportunities in scientific and technological fields	1	2	3	4	5
j. Discuss controversial inventions and technologies	1	2	3	4	5

18. For biology teachers: Have you taught or reviewed the following topics in this Biology class during this year? If you have reviewed and taught an item as new content, mark #3 only.

(Circle One Number on Each Line)

	No, but it was was taught previously	Yes, but I reviewed it only	Yes, but I taught it as new content	No, but I will teach or review it later this year	No, topic is beyond the scope of this course
a. Cell structure and function	1	2	3	4	5
b. Genetics	1	2	3	4	5
c. Diversity of life	1	2	3	4	5
d. Metabolism and regulation of the organism	1	2	3	4	5
e. Behavior of the organism	1	2	3	4	5
f. Reproduction and development of the organism	1	2	3	4	5
g. Human biology	1	2	3	4	5
h. Evolution	1	2	3	4	5
i. Ecology	1	2	3	4	5

19. For physics teachers: Have you taught or reviewed the following topics in this Physics class during this year? If you have reviewed and taught an item as new content, mark #3 only.

(Circle One Number on Each Line)

	No, but it was was taught previously	Yes, but I reviewed it only	Yes, but I taught it as new content	No, but I will teach or review it later this year	No, topic is beyond the scope of this course
a. Forms and sources of energy	1	2	3	4	5
b. Forces, time, motion	1	2	3	4	5
c. Molecular/nuclear physics	1	2	3	4	5
d. Energy/matter transformations	1	2	3	4	5
e. Sound and vibrations	1	2	3	4	5
f. Light	1	2	3	4	5
g. Electricity and magnetism	1	2	3	4	5
h. Solids/fluids/gases	1	2	3	4	5

ERIC

Teacher Background and Experience

1. **What is your sex?**

   Male.................................. 1

   Female 2

2. **Which best describes you?**

   Asian or Pacific Islander..................................1
   Hispanic, regardless of race 2
   Black, not of Hispanic origin 3
   White, not of Hispanic origin 4
   American Indian or Alaskan Native 5

3. **What is the year of your birth?**

   (Last 2 digits): ☐☐

4. **Counting this year, how many years in total have you taught at either the elementary or secondary level?**

   K-6: ☐☐

   7-12: ☐☐

5. **Counting this year, how many years in total have you taught in this school?**

   Years: ☐☐

ERIC

6. **What academic degree(s) do you hold?**

(Circle All That Apply)

No degree	0	—> SKIP TO Q8
Associate degree	1	—> SKIP TO Q8, if only degree
Bachelor's	2	
Master's	3	
Education specialist or professional diploma at least one year of work beyond master's level	4	
Doctorate	5	
First professional degree (e.g., M.D., D.D.S.)	6	

7. **What were your major and minor fields of study for your bachelor's degree?**

(Circle All That Apply)

	Major	Minor
a. Education	1	1
b. Mathematics	2	2
c. Natural/physical sciences	3	3
d. Life/biological sciences	4	4
e. Computer science	5	5
f. Foreign language	6	6
g. English	7	7
h. History (or social science)	8	8
i. Other	9	9

8. **Circle the number beside any of the following subjects which you have taught this year.**

(Circle All That Apply)

MATHEMATICS

General Math	01
Pre-Algebra	02
Algebra I	03
Algebra II	04
Geometry	05
Trigonometry	06
Pre-Calculus	07
Calculus	08
Consumer/Business Math	09
AP Calculus	10
Other Math	11

SCIENCE

General Science	12
General Physical Science	13
Earth Science	14
Principles of Technology	15
Biology	16
Chemistry	17
Physics	18
AP Science	19
Other Science	20

OTHER

Computer Science	21
Other non-math, non-science course	22

Please describe ____________________

Date completed: ☐☐ / ☐☐ / ☐☐
MO DAY YR

Thank you for your assistance.

Please return this survey in the same envelope with your first week's instructional materials.

ERIC Full Text Provided by ERIC

00000

VALIDATING NATIONAL CURRICULUM INDICATORS

MATHEMATICS TEACHER SURVEY

00000

As part of the larger study to examine different ways of measuring curriculum trends in schools, this questionnaire asks you to report on the goals, content, and instructional activities in the class for which you have been providing us with your instructional materials. Specifically, it asks about the curriculum content covered, the teaching strategies and instructional practices used, and your goals, objectives and general beliefs about the way mathematics should be taught to this class The information you provide, along with other data already collected, is intended to describe students' educational experiences. Also, because this study will inform future efforts, space is provided at the end of the questionnaire for your comments on any problems or recommendations.

Please MARK YOUR RESPONSES DIRECTLY ON THE QUESTIONNAIRE. Place it in the envelope with your instructional materials for this week, and return it to RAND.

THANK YOU FOR YOUR CONTRIBUTION TO THE STUDY.

ERIC

Class Characteristics

Please provide the following information about the specific class listed below:

Designated class: ______________________________________ 00000

1. How many students are in this class? ________ Total 00000

   ________ Females 00000

   ________ Males 00000

2. How many of the students in this class are in the following grade levels? (Sum should equal total number of students given above.)

   a. 9th grade ________ 00000

   b. 10th grade ________ 00000

   c. 11th grade ________ 00000

   d. 12th grade ________ 00000

3. Which of the following best describes the achievement level of the students in this class in comparison to the average student in this school? (Circle one.)

   This class consists primarily of students with:

Higher achievement levels	1	00000
Average achievement levels	2	00000
Lower achievement levels	3	00000
Widely differing achievement levels	4	00000

4. How many of the students in this class are of limited or non-English speaking ability?

   ________ 00000

Form 1 9/17/92

ERIC

5. **How many of the students in this class are members of the following ethnic/racial groups? (Sum should equal total given above in question 1.)**

a. American Indian or Alaskan Native ________ 00000

b. Asian or Pacific Islander ________ 00000

c. Hispanic, regardless of race ________ 00000

d. Black (not of Hispanic origin) ________ 00000

e. White (not of Hispanic origin) ________ 00000

f. Other (specify) ______________ ________ 00000

6. **How many students in this class are likely to do the following in the future? (Sum should equal total given above in question 1.)**

a. Attend a 4-year college ________ 00000

b. Attend a 2-year college/technical school ________ 00000

c. End formal education with high school ________ 00000

d. Not graduate from high school ________ 00000

Curriculum Coverage

Please answer the following questions about the content you taught this class.

7. What was the primary text used in this class?

Title: ____ 00000

8. What chapters do you plan to cover by the end of this semester?

Chapters: ____ 00000

How closely did you follow the text? (Describe your use of the text below.)

____ 00000

____

9. What additional chapters do you plan to cover over the course of this year?

Chapters: ____ 00000

You will find a list of topics on this page and the next 2 pages. Please respond to the following questions for each of the topics listed.

10. Have you taught or reviewed the following topics during this year in this class? (Circle your response.)

    1 = No, but it was taught previously.

    2 = Yes, but I reviewed it only.

    3 = Yes, I taught it as new content (includes new topics which will be reviewed later).

    4 = Not yet, but I will teach or review it later this school year.

    5 = No, topic is beyond the scope of this course or not in the school curriculum.

11. Indicate the approximate number of periods devoted to each topic below. If you focus on a topic for 10 or 15 minutes on a given day, count that as a period. If you will teach or review a topic later this year, indicate the number of periods you anticipate spending on the topic. (Circle your response.)

    1 = None (zero)

    2 = One or two periods

    3 = Three to five periods

    4 = Six to ten periods

    5 = More than two weeks but less than one month (11 to 20 class periods)

    6 = One month or more (more than 20 periods)

BEST COPY AVAILABLE

Topics:	10. Taught or reviewed? No, Taught previously	Yes, Reviewed Only	Yes, New content	Not yet, Later in year	No, Beyond course	11. Periods on topic? 0 Periods	1-2 Periods	3-5 Periods	6-10 Periods	11-20 Periods	> 20 Periods
a. Patterns and functions	1	2	3	4	5	1	2	3	4	5	6
b. Estimation	1	2	3	4	5	1	2	3	4	5	6
c. Proportional reasoning	1	2	3	4	5	1	2	3	4	5	6
d. Proofs	1	2	3	4	5	1	2	3	4	5	6
e. Tables and charts	1	2	3	4	5	1	2	3	4	5	6
f. Graphing	1	2	3	4	5	1	2	3	4	5	6
g. Math modeling	1	2	3	4	5	1	2	3	4	5	6
* h. Ratios, proportions, and percents	1	2	3	4	5	1	2	3	4	5	6
* i. Conversions among fractions, decimals and percents	1	2	3	4	5	1	2	3	4	5	6
* j. Laws of exponents	1	2	3	4	5	1	2	3	4	5	6
* k. Square roots	1	2	3	4	5	1	2	3	4	5	6
l. Polynomials	1	2	3	4	5	1	2	3	4	5	6
* m. Linear equations	1	2	3	4	5	1	2	3	4	5	6
n. Slope	1	2	3	4	5	1	2	3	4	5	6
* o. Writing equations for lines	1	2	3	4	5	1	2	3	4	5	6
* p. Inequalities	1	2	3	4	5	1	2	3	4	5	6
q. Quadratic equations	1	2	3	4	5	1	2	3	4	5	6
* r. Applications of measurement formulas (e.g. area, volume)	1	2	3	4	5	1	2	3	4	5	6
s. Properties of geometric figures	1	2	3	4	5	1	2	3	4	5	6
* t. Pythagorean Theorem	1	2	3	4	5	1	2	3	4	5	6

		10. Taught or reviewed?					11. Periods on topic?					
		No, Taught previously	Yes, Reviewed Only	Yes, New content	Not yet, Later in year	No, Beyond course	0 Periods	1-2 Periods	3-5 Periods	6-10 Periods	11-20 Periods	> 20 Periods
u.	Coordinate geometry	1	2	3	4	5	1	2	3	4	5	6
v.	Probability	1	2	3	4	5	1	2	3	4	5	6
w.	Statistics	1	2	3	4	5	1	2	3	4	5	6
* x.	Distance, rate, time problems	1	2	3	4	5	1	2	3	4	5	6
y.	Growth and decay	1	2	3	4	5	1	2	3	4	5	6
** z.	Transformational geometry	1	2	3	4	5	1	2	3	4	5	6
** aa.	Logarithms	1	2	3	4	5	1	2	3	4	5	6
** bb.	Conic sections	1	2	3	4	5	1	2	3	4	5	6
** cc.	Trigonometry	1	2	3	4	5	1	2	3	4	5	6
** dd.	Polar coordinates	1	2	3	4	5	1	2	3	4	5	6
** ee.	Sequences	1	2	3	4	5	1	2	3	4	5	6
** ff.	Complex numbers	1	2	3	4	5	1	2	3	4	5	6
** gg.	Vectors	1	2	3	4	5	1	2	3	4	5	6
** hh.	Matrices and matrix operations	1	2	3	4	5	1	2	3	4	5	6
** ii.	Calculus	1	2	3	4	5	1	2	3	4	5	6
** jj.	Limits	1	2	3	4	5	1	2	3	4	5	6
** kk.	Integration	1	2	3	4	5	1	2	3	4	5	6
** ll.	Fundamental counting principle, permutations, combinations	1	2	3	4	5	1	2	3	4	5	6
** mm.	Measures of dispersion (range variance, standard deviation, etc.)	1	2	3	4	5	1	2	3	4	5	6
** nn.	Discrete math (e.g., Euler circuits, directed graphs, trees)	1	2	3	4	5	1	2	3	4	5	6

* indicates topic in Form I only.
** indicates topic in Form II only.

ERIC

12. **For each item below, please indicate the types of <u>student understanding</u> you expect from the majority of this class by the end of the course. (Circle the highest number that applies.)**

1 = Recognizes/knows the rule or principle
2 = When given the rule or principle, is able to use it
3 = Knows when and how to apply the rule or principle
4 = Can both apply the rule or principle and explain why it works as it does
5 = Not applicable—rule or principle beyond the scope of this class

a. Division by zero is not allowed: $\frac{a}{0}$ is undefined for all numbers a	1	2	3	4	5	00000
b. In a plane, the sum of the angle measures in any triangle is 180	1	2	3	4	5	00000
c. The area of a triangle: $A = \frac{1}{2}bh$	1	2	3	4	5	00000
d. The Pythagorean Theorem	1	2	3	4	5	00000
e. The slope of a vertical line is undefined	1	2	3	4	5	00000
f. The distance formula: $d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$	1	2	3	4	5	00000
h. If $\frac{a}{b} = \frac{c}{d}$, then $ad = bc$	1	2	3	4	5	00000
h. $(a+b)^2 = a^2 + 2ab + b^2$	1	2	3	4	5	00000
i. The product rule for exponents: $a^m \cdot a^n = a^{m+n}$	1	2	3	4	5	00000
j. The square root of a negative number is not a real number	1	2	3	4	5	00000
k. The log of a negative number is not defined	1	2	3	4	5	00000
l. A continuous function need not be differentiable	1	2	3	4	5	00000

Instructional Practices

Please answer the following questions about the organization, teaching strategies and instructional practices you used with this class.

13. How often do you use each of the following instructional strategies with this class? (The strategy need not take the entire class period.)

	Almost every day	Once or twice a week	Once or twice a month	Once or twice a semester	Never	
a. Lecture	1	2	3	4	5	00000
b. Have students respond orally to questions on subject matter	1	2	3	4	5	00000
c. Have student-led whole group discussions	1	2	3	4	5	00000
d. Have teacher-led whole group discussions	1	2	3	4	5	00000
e. Correct and/or review homework in class	1	2	3	4	5	00000
f. Demonstrate working an exercise at the board	1	2	3	4	5	00000
g. Have students work exercises at the board	1	2	3	4	5	00000
h. Have students work individually on written assignments or worksheets in class	1	2	3	4	5	00000
i. Have students give oral reports	1	2	3	4	5	00000
j. Administer a test (full period)	1	2	3	4	5	00000
k. Administer a quiz	1	2	3	4	5	00000
l. Use manipulatives (e.g., conic section models) to demonstrate a concept	1	2	3	4	5	00000

(Continued on next page.)

	Almost every day	Once or twice a week	Once or twice a month	Once or twice a semester	Never	
m. Discuss career opportunities in mathematics	1	2	3	4	5	00000
n. Have small groups work on problems to find a joint solution	1	2	3	4	5	00000
o. Have whole class discuss solutions developed in small groups	1	2	3	4	5	00000
p. Have students practice or drill on computational skills	1	2	3	4	5	00000
q. Have students work on problems for which there is no obvious method of solution	1	2	3	4	5	00000
r. Have students represent and analyze relationships using tables and graphs	1	2	3	4	5	00000
s. Have students use calculators to solve exercises or problems	1	2	3	4	5	00000
t. Have students use computers to solve exercises or problems	1	2	3	4	5	00000
u. Have students respond to questions or assignments that require writing at least a paragraph	1	2	3	4	5	00000
v. Have students keep a mathematics journal	1	2	3	4	5	00000
w. Have students read textbooks or supplementary materials	1	2	3	4	5	00000
x. Have students work with manipulatives	1	2	3	4	5	00000
y. Have students work on next day's homework in class	1	2	3	4	5	00000
z. Summarize main points of today's lesson	1	2	3	4	5	00000
aa. Have students work on projects in class	1	2	3	4	5	00000

14. Indicate what percent of class time is spent in a typical week doing each of the following with this class. (Circle one on each line. The total need not sum to 100%.)

	Percent None	< 10	10-24	25-49	50-74	75-100	
a. Providing instruction to the *class* as a whole	1	2	3	4	5	6	00000
b. Providing instruction to *small groups* of students	1	2	3	4	5	6	00000
c. Providing instruction to *individual students*	1	2	3	4	5	6	00000
d. Maintaining order/disciplining students	1	2	3	4	5	6	00000
e. Administering tests or quizzes	1	2	3	4	5	6	00000
f. Performing routine administrative tasks (e.g., taking attendance, making announcements, etc.)	1	2	3	4	5	6	00000
g. Conducting lab periods	1	2	3	4	5	6	00000

Evaluation and Grading Practices

15. On the tests, quizzes, and exams you administer to this class, about what percent of the items are of the following types? (Total should equal 100% in each column.)

	Tests and Quizzes	Final Exam	
a. Multiple-choice	________%	________%	00000
b. Short-answer	________%	________%	00000
c. Essay	________%	________%	00000
d. Open-ended problems (i.e., where students generate their own solutions)	________%	________%	00000
e. Other (specify) ____________________	________%	________%	00000

16. On the tests and quizzes you administer to this class, about what percent of the items are of the following types? (Total need not sum to 100%.)

a. Items that require students to recognize or recall definitions or concepts	________%	00000
b. Items that require the use of algorithms to solve problems	________%	00000
c. Items that require students to describe how to solve problems	________%	00000
d. Items that require students to explain their reasoning	________%	00000
e. Items that require the application of concepts or principles to different or unfamiliar situations	________%	00000
h. Items that require a critique or analysis of a suggested solution to a problem	________%	00000
i. Other (specify) ________________________	________%	00000

17. On the tests and quizzes you administer to this class, about what percent of the items are of the following types? (Total need not sum to 100%.)

a. Exercises or problems that are minor variations of homework or class exercises or problems	________%	00000
b. Exercises or problems with more than one possible answer	________%	00000
c. Exercises or problems with more than one possible approach	________%	00000
d. Exercises or problems that require more than one step to reach a solution	________%	00000
e. Items that require the use of tabular or graphical data	________%	00000

18. What will be the approximate distribution of final student grades in this class? (Total should equal number of students in the class.)

A's ____________	00000
B's ____________	00000
C's ____________	00000
D's ____________	00000
F's ____________	00000

Homework Policies and Practices

19. Approximately how much homework do you typically assign each day to this class?

________________ ___________ minutes 00000

20. How often do you do each of the following with homework assignments?

	Never	Some of the time	Most of the time	All of the time	
a. Keep records of who did or who turned in the assignment	1	2	3	4	00000
b. Return assignments with grades or corrections	1	2	3	4	00000
c. Discuss the completed assignment in class	1	2	3	4	00000

21. How frequently do you assign each of the following types of homework?

	Almost every day	Once or twice a week	Once or twice a month	Once or twice a semester	Never	
a. Reading the text or supplementary materials	1	2	3	4	5	00000
b. Doing exercises or problems from the text	1	2	3	4	5	00000
c. Doing exercises or problems from worksheets	1	2	3	4	5	00000
d. Writing definitions of concepts	1	2	3	4	5	00000
e. Applying concepts or principles to different or unfamiliar situations	1	2	3	4	5	00000
f. Solving problems for which there is no obvious method of solution	1	2	3	4	5	00000
g. Gathering data, conducting experiments, working on projects	1	2	3	4	5	00000
h. Preparing oral reports	1	2	3	4	5	00000
i. Preparing written reports	1	2	3	4	5	00000
j. Extending results established in class (e.g., deriving or proving new results)	1	2	3	4	5	00000
k. Keeping a journal	1	2	3	4	5	00000
l. Solving applied problems (e.g., finding the amount of water needed to fill a pool)	1	2	3	4	5	00000
m. Explaining newspaper/magazine articles	1	2	3	4	5	00000

ERIC

Materials, Equipment and Technology

22. How frequently do you use the materials and equipment listed below with this class?

	Almost every day	Once or twice a week	Once or twice a month	Once or twice a semester	Never	
a. Graph paper	1	2	3	4	5	00000
b. Protractors, rulers, or compasses	1	2	3	4	5	00000
c. A-V equipment (e.g., film projector, VCR, cassette, TV)	1	2	3	4	5	00000
d. Overhead projector	1	2	3	4	5	00000
f. Four-function calculator	1	2	3	4	5	00000
g. Scientific calculator	1	2	3	4	5	00000
h. Graphing calculator	1	2	3	4	5	00000
i. Other (specify)	1	2	3	4	5	00000

____

Goals, Objectives and Teacher Beliefs

23. How much emphasis do you give to each of the following objectives in this class?

	Emphasis				
	No	Minor	Moderate	Major	
a. Understanding the nature of proof	1	2	3	4	00000
b. Memorizing facts, rules and steps	1	2	3	4	00000
c. Learning to represent problem structures in multiple ways (e.g., graphically, algebraically, numerically)	1	2	3	4	00000
d. Integrating different branches of mathematics (e.g., algebra, geometry) into a unified framework	1	2	3	4	00000
e. Conceiving and analyzing the effectiveness of different approaches to problem solving	1	2	3	4	00000
f. Performing calculations with speed and accuracy	1	2	3	4	00000
g. Showing the importance of math in daily life	1	2	3	4	00000
h. Solving equations	1	2	3	4	00000
i. Raising questions and formulating conjectures	1	2	3	4	00000
j. Increasing students' interest in math	1	2	3	4	00000
k. Integrating math with other subjects	1	2	3	4	00000
l. Finding examples and counterexamples	1	2	3	4	00000
m. Judging the validity of arguments	1	2	3	4	00000
n. Discovering generalizations	1	2	3	4	00000
o. Representing and analyzing relationships using tables, charts and graphs	1	2	3	4	00000
p. Applying mathematical models to real-world phenomena	1	2	3	4	00000
q. Writing about mathematical ideas	1	2	3	4	00000
r. Designing a study or experiment	1	2	3	4	00000
s. Writing equations to represent relationships	1	2	3	4	00000
t. Solving problems for which there is no obvious method of solution	1	2	3	4	00000

ERIC

24. Indicate the degree to which you emphasized the following strategies with this class.

	Emphasis				
	No	Minor	Moderate	Major	
a. Students received a good deal of practice to become competent at mathematics.	1	2	3	4	00000
b. I routinely justified the mathematical principles and procedures used.	1	2	3	4	00000
c. I corrected student errors immediately.	1	2	3	4	00000
d. Students were provided frequent opportunities to discover mathematical ideas for themselves.	1	2	3	4	00000
e. I gave step-by-step directions for applying algorithms and procedures.	1	2	3	4	00000
f. Students were provided opportunities to apply mathematics to real-world situations.	1	2	3	4	00000
g. Students developed their own methods of solving math problems.	1	2	3	4	00000
h. Students were frequently expected to discover generalizations and principles on their own.	1	2	3	4	00000
i. Students learned to solve problems in different ways.	1	2	3	4	00000
j. Students were required to memorize and apply rules.	1	2	3	4	00000
k. Students learned there is usually a rule to apply when solving a math problem.	1	2	3	4	00000
l. Students received step-by-step directions to aid in solving problems.	1	2	3	4	00000

25. There are a variety of ways in which teachers describe their role in helping their students learn mathematics. Statements A through D represent several possibilities. Please read these statements, then answer the question below about your role.

A: "I mainly see my role as a facilitator. I try to provide opportunities and resources for my students to discover or construct mathematical concepts for themselves."

B: "I think I need to provide more guidance than that. Although I provide opportunities for them to discover concepts, I also try to lead my students to figure things out by asking pointed questions without telling them the answers."

C: "I emphasize student discussion of math in my classroom. We talk about concepts and problems together, exploring the meaning and evaluating the reasoning that underlies different strategies. My role is to initiate and guide these discussions."

D: "That's all nice, but my students really won't learn math unless you go over the material in a detailed and structured way. I think it's my job to explain, to show students how to do the work, and to give them practice doing it."

Which statement best typifies your conception of your role in helping students in this class learn math? (Place an X on the continuum below to indicate your role.)

A B C D

00000

26. Below are two pairs of statements. Each pair represents opposite ends of a continuum in curriculum approaches. After reading a pair of statements, place an X on the line between that pair indicating where you would place your approach with this class.

<u>Pair 1:</u> My primary goal is to help students

A: learn mathematical terms, master computational skills and solve word problems

B: achieve a deeper conceptual understanding of mathematics

A ________________________________ B

00000

<u>Pair 2:</u> In this mathematics class, I aim for

A: in-depth study of selected topics and issues, even if it means sacrificing coverage

B: comprehensive coverage even if it means sacrificing in-depth study

A ________________________________ B

00000

ERIC

The following questions concern the questionnaire itself. Please provide this information so that we might improve the questionnaire for future use.

27. Were any of the questions confusing or unclear?

_______ No

_______ Yes If Yes, please list the question number and describe the source of confusion.

Number	Source of Confusion
_______	____
_______	____
_______	____
_______	____
_______	____
_______	____
_______	____
_______	____
_______	____

28. Use the space below to describe any other problems or make any recommendations about the questionnaire.

____

____

____

____

____

____

____

____

Daily Log

Name____________________ Course____________________

Date____________________ School____________________

1. List the content covered in this class period by briefly describing it or providing examples.

TOPICS

2. What modes of instruction did you use? (Check all that apply.)

Lecture to entire class ________
Demonstrate an exercise at the board ________
Use manipulatives or audio-visual materials to demonstrate a concept ________
Demonstrate an experiment ________
Lead question and answer session ________
Work with small groups ________
Work with individual students ________
Correct or review homework ________
Other (please specify)____________________ ________

3. What activities did students engage in during this period? (Check all that apply.)

Listen and take notes ________
Work exercises at board ________
Work individually on written assignments or worksheets ________
Work with other students ________
Work with manipulatives ________
Use calculators ________
Respond to questions ________
Discuss topics from lesson ________
Work on next day's homework ________
Work on computer ________
Conduct lab experiment ________
Write lab report ________
Other (please specify)____________________ ________

Comments:____

ERIC Full Text Provided by ERIC